About positive, energy conservative and equilibrium state preserving schemes for the isotropic Fokker-Planck-Landau equation

Christophe Buet, Kim-Claire Le Thanh

CEA-R-6121

ISSN 0429 - 3460

COMMISSARIAT À L'ÉNERGIE ATOMIQUE

ABOUT POSITIVE, ENERGY CONSERVATIVE AND EQUILIBRIUM STATE PRESERVING SCHEMES FOR THE ISOTROPIC FOKKER-PLANCK-LANDAU EQUATION

par
Christophe BUET
et Kim-Claire LE THANH

CEA/DAM ILE-DE-FRANCE

DÉPARTEMENT SCIENCES DE LA SIMULATION ET DE L'INFORMATION

SERVICE NUMÉRIQUE ENVIRONNEMENT ET CONSTANTES

DÉPARTEMENT DE PHYSIQUE THÉORIQUE ET APPLIQUÉE

SERVICE PHYSIQUE DES PLASMA ET ELECTROMAGNÉTISME

DIRECTION DES SYSTÈMES D'INFORMATION

CEA / SACLAY 91191 GIF-SUR-YVETTE CEDEX FRANCE

CEA/DAM ILE-DE-FRANCE

2006

RAPPORT CEA-R-6121

- Rapport CEA-R-6121 -

CEA/DAM Ile de France
Département Sciences de la Simulation et de L'Information
Service Numérique Environnement et Constantes
&
Département de Physique Théorique et Appliquée
Service Physique des Plasma et Électromagnétisme

ABOUT POSITIVE, ENERGY CONSERVATIVE AND EQUILIBRIUM STATE PRESERVING SCHEMES FOR THE ISOTROPIC FOKKER-PLANCK-LANDAU EQUATION

par

Christophe BUET
&
Kim-Claire LE THANH

- Septembre 2006 -

RAPPORT CEA-R-6121 – Christophe BUET, Kim-Claire LE THANH

«Sur les schémas positifs, conservant l'énergie et les états d'équilibre pour l'équation de Fokker-Planck-Landau isotrope»

Résumé - Dans ce rapport on s'intéresse à la discrétisation des termes de collisions de l'équation de Fokker-Planck-Landau. En particulier on étudie les collisions entre des électrons isotropisés par un bain de particules lourdes, par exemple des ions. La discussion porte sur les schémas positifs qui conservent la masse, l'énergie et les états d'équilibre maxwelliens. En premier lieu on analyse en détails le schéma de Chang et Cooper pour ce terme de collisions non linéaires : construction du schéma, positivité et propriétés de conservation. On montre que certaines variantes de ce schéma, dérivées de l'écriture de l'opérateur de Fokker-Planck sous la forme Rosenbluth, peuvent ne pas être positives ou ne pas conserver l'énergie. Nous présentons aussi une nouvelle version du schéma de Chang et Cooper, écrite à partir de la forme Landau, qui possède toutes les bonnes propriétés. Pour terminer, on propose deux autres nouveaux schémas, eux aussi dérivés de la forme Landau, qui montrent que le schéma de Chang et Cooper n'est pas le seul à posséder de bonnes propriétés.
Pour tous ces schémas, nous analysons clairement les propriétés de conservation de la densité et de l'énergie, la positivité des solutions et la conservation des états d'équilibre. On s'intéresse également aux cas où les collisions sont maxwelliennes ou coulombiennes.

2006 – Commissariat à l'Énergie Atomique – France

RAPPORT CEA-R-6121 – Christophe BUET, Kim-Claire LE THANH

« About positive, energy conservation and equilibrium state preserving schemes for the isotropic Fokker-Planck-Landau equation»

Abstract **-** The aim of this paper is to describe the discretization of the Fokker-Planck-Landau (FPL) collision term in the isotropic case which models the self collision for the electrons when they are totally isotropized by heavy particles background such as ions. The discussion focus on schemes which could preserve positivity, mass energy and Maxwellian equilibrium. First, we analyze in detail the popular Chang and Cooper method for this non-linear collision term : derivation, conservation and positivity properties. We show that some variants of this method, based on the drift-diffusion form of the FPL operator, could not be positive or could not conserve the energy. We present a new variant of the Chang and Cooper method derived from the Landau form that is both positive and conservative. We also propose two new alternatives and simpler schemes for the FPL operator which show that the Chang and Cooper is not the only way to construct positive, energy conservative and equilibrium state preserving schemes for this operator. For all these schemes, we explain clearly the properties of conservation of the density and the energy, the positivy of the solution and the conservation of equilibrium states, or their lack. The case of Maxwellian and Coulombian potentials are emphasized.

2006 – Commissariat à l'Énergie Atomique – France

About positive, energy conservative and equilibrium state preserving schemes for the isotropic Fokker-Planck-Landau equation

Christophe Buet
Département des Sciences de la Simulation et de l'Information,
CEA-DIF,
91680, Bruyères le Châtel, BP 12, France

Kim-Claire Le Thanh
Département de Physique Théorique et Appliquée,
CEA-DIF,
91680, Bruyères le Châtel, BP 12, France

Abstract

The aim of this paper is to describe the discretization of the Fokker-Planck-Landau (FPL) collision term in the isotropic case which models the self collision for the electrons when they are totally isotropized by heavy particles background such as ions. The discussion focus on schemes which could preserve positivity, mass, energy and Maxwellian equilibrium. First, we analyze in detail the popular Chang and Cooper method for this non-linear collision term: derivation, conservation and positivity properties. We show that some variants of this method, based on the drift-diffusion form of the FPL operator, could not be positive or could not conserve the energy. We present a new variant of the Chang and Cooper method derived from the Landau form that is both positive and conservative. We also propose two new alternatives and simpler schemes for the FPL operator which show that the Chang and Cooper method is not the only way to construct positive, energy conservative and equilibrium state preserving schemes for this operator.

For all these schemes, we explain clearly the properties of conservation of the density and the energy, the positivity of the solution and the conservation of the equilibrium states, or their lack. The case of Maxwellian and Coulombian potentials are emphasized.

Contents

1 Introduction

The Fokker-Planck-Landau equation is commonly used in plasma physics when studying kinetic effects between charged particles under coulomb interaction.
The isotropic Fokker-Planck-Landau operator is generally used in the modelling of inertial controlled fusion. More precisely, it is used to describe the electronic energy transport phenomena in laser produced plasma. In some conditions, it is well known that the fluid theory, for which the hydrodynamic equations are closed using the law for the thermal fluxes proposed by Spitzer-Harm [29], is not valid [21, 22]. A more accurate solution is to use a model based on the expansion of the FPL operator in spherical harmonics and to retain the two first terms, and the isotropic FPL operator is the leading order term [21, 22, 23]. There are also applications in the astrophysical field where the FPLE is used for star's clusters modelling [11, 12].
The most popular scheme for Fokker-Planck type equations is the Chang and Cooper method [10]. This method was originally devoted to linear Fokker-Planck equations and it was shown in [10] that in this case this method is positive and preserves the equilibrium states. On the contrary a paper of Larsen *et al.* [25] shows that this method applying for nonlinear Fokker-Planck equations could produce non positive solutions. For all that, this method is used for the isotropic Landau equation by Langdon [24], in the SPARK code by Epperlein [22], by Kingham and Bell [23] in the code IMPACT or by Cohn [11] in the astrophysical context. To our knowledge, there is no rigorous proof of the positivity and the conservation of the energy of this scheme when it is applied to isotropic Landau equation. For the isotropic Landau equation the Chang and Cooper scheme is derived from the Rosenbluth form of the equation which is not the most powerful form to check energy conservation and could lead to mistake at boundary.
There exist also conservative, positive and entropy schemes for this equation [1, 4, 5, 17, 26, 27] based on the "Log" weak symmetric form of the Landau equation, but these methods suffer from some limitations due to the use of Logarithm of the distribution function or, when not using "Log", they can only be defined for uniform mesh in the velocity space.

Thus it would be interesting to clarify the situation about positive, conservative and equilibrium states preserving velocity discretization of the isotropic Landau equation. This is the aim of this work.
In the most important part of this paper we focus our attention on the Chang and Cooper method. By starting from the weak symmetrized form of the isotropic Landau operator we construct two variants of the Chang and Cooper method. A new one, called $\mathcal{S}_1$, and an other one, called $\mathcal{S}_2$, that is described by Langdon *et al.* [15, 24] in the Coulombian case but constructed directly from the Rosenbluth form of the operator. These two schemes only differ in the boundary condition. We prove that these two variants conserve the energy but we show that the variant which fits with the Langdon's scheme is not positive.
Moreover we make also some remarks about simplifications of these schemes which could lead to the loss of the positivity or the loss of the energy conservation. These remarks would be useful for the time discretization, especially for time implicit discretization. We clarify also the boundary conditions at the end of the domain. As we will see this leads to the conclusion that, due to bad boundary conditions, the scheme used by Epperlein [22] or by Kingham and Bell [23], and also obtained using the Rosenbluth form, is not conservative in energy (because in collision operator we can't reverse the role of test particles and target particles) but it is positive.
In other hand we also show that the Chang and Cooper method is not the only way to obtain positive, conservative and equilibrium preserving states even on non-uniform meshes. We propose two new other schemes that share the same properties with the Chang and Cooper method. One of them, called $\mathcal{S}_3$, is based on the work of Larsen *et al.* [25]. The other, called $\mathcal{S}_4$, is based on the "Log" form of the equation (it is a new version of the scheme described in [17] applied to the isotropic FPL equation). Thus it is also an entropy decaying scheme. In the case of uniform meshes these two schemes are nothing but the conservative positive and entropy scheme of Berezin *et al.* [1] and studied in detail by one of the authors [4, 5].
This paper is organized as follow: in the first part, we recall the continuous FPL equation in the homogeneous and isotropic case and its properties. We recall also the different forms of the equation in the case of Maxwellian or Coulombian potentials.
In a second part we present the two Chang and Cooper type schemes and the two non Chang and Cooper ones. We clearly show all the properties of these schemes: energy conservation, equilibrium states and positivity. We also indicate for each of them, the simplifications in the case of a uniform mesh, or in the maxwellian case. Some comments are also made about the advantages or disadvantages of each of these four schemes especially for the implementation.

2 Fokker-Planck-Landau equation

We present the homogeneous non-linear Fokker-Planck-Landau equation (FPL equation) in the isotropic case where the distribution function $f(\vec{x},\vec{v},t)$ depends only on the modulus of the velocity $v = \| \vec{v} \|$ and on the time t, in other words $f(\vec{x},\vec{v},t) = f(v,t)$. We consider f as a function of $\varepsilon = v^2$ where ε is the energy variable. In this case, Fokker-Planck-Landau equation can be written:

$$\frac{\partial f(\varepsilon,t)}{\partial t} = Q(f)(\varepsilon) = \frac{1}{\sqrt{\varepsilon}}\frac{\partial}{\partial \varepsilon}\int_0^\infty g(\varepsilon,\varepsilon')\left(f(\varepsilon')\frac{\partial f(\varepsilon)}{\partial \varepsilon} - f(\varepsilon)\frac{\partial f(\varepsilon')}{\partial \varepsilon'}\right)d\varepsilon', \qquad (2.1)$$

where $g(\varepsilon,\varepsilon')$ is positive, symmetric and increasing ($g(\varepsilon,\varepsilon') = \min(\varepsilon^{\frac{3}{2}},\varepsilon'^{\frac{3}{2}})$ for Coulombian

interactions and $g(\varepsilon,\varepsilon') = \varepsilon^{\frac{3}{2}}\varepsilon'^{\frac{3}{2}}$ for Maxwellian interactions) and $Q(f)$ is the so-called FPL collision operator. This operator can be written in the following weak form (let $\phi(\varepsilon)$ be any test (smooth and decaying) time independent function)

$$\int_0^\infty \frac{\partial f(\varepsilon,t)}{\partial t}\,\phi(\varepsilon)\sqrt{\varepsilon}d\varepsilon \;=\; \int_0^\infty \phi(\varepsilon)\frac{\partial}{\partial\varepsilon}\left(\int_0^\infty g(\varepsilon,\varepsilon')(f(\varepsilon')\frac{\partial f(\varepsilon)}{\partial\varepsilon} \;-\; f(\varepsilon)\frac{\partial f(\varepsilon')}{\partial\varepsilon'})\,d\varepsilon'\right)d\varepsilon\,, \tag{2.2}$$

and by integrating (2.2) by parts

$$\int_0^\infty \frac{\partial f(\varepsilon,t)}{\partial t}\,\phi(\varepsilon)\sqrt{\varepsilon}d\varepsilon \;=\; -\frac{1}{2}\int_0^\infty\int_0^\infty (\frac{\partial\phi(\varepsilon)}{\partial\varepsilon}-\frac{\partial\phi(\varepsilon')}{\partial\varepsilon'})\,g(\varepsilon,\varepsilon')\,(f(\varepsilon')\frac{\partial f(\varepsilon)}{\partial\varepsilon} - f(\varepsilon)\frac{\partial f(\varepsilon')}{\partial\varepsilon'})\,d\varepsilon'\,d\varepsilon\,. \tag{2.3}$$

Note that the FPL equation can be equivalently written in the so-called "Log" weak form

$$\int_0^\infty \frac{\partial f(\varepsilon,t)}{\partial t}\,\phi(\varepsilon)\sqrt{\varepsilon}d\varepsilon = -\frac{1}{2}\int_0^\infty\int_0^\infty (\frac{\partial\phi(\varepsilon)}{\partial\varepsilon}-\frac{\partial\phi(\varepsilon')}{\partial\varepsilon'})g(\varepsilon,\varepsilon')f(\varepsilon')f(\varepsilon)(\frac{\partial\log f(\varepsilon)}{\partial\varepsilon}-\frac{\partial\log f(\varepsilon')}{\partial\varepsilon'})\,d\varepsilon'd\varepsilon\,. \tag{2.4}$$

2.1 Properties

Let us recall the most important properties of the problem (2.1)

1. This operator satisfies the conservation of mass (respectively energy) by choosing $\phi(\varepsilon) = 1$ (respectively $\phi(\varepsilon) = \varepsilon$) in (2.3)

$$\boldsymbol{\rho} \;=\; \int_0^\infty f(\varepsilon,t)\sqrt{\varepsilon}d\varepsilon\,, \tag{2.5}$$

$$\boldsymbol{\rho}\mathbf{E} \;=\; \int_0^\infty f(\varepsilon,t)\varepsilon^{\frac{3}{2}}d\varepsilon\,. \tag{2.6}$$

   Note that the conservation properties are a consequence of the symmetry property (between ε and ε') of the collision operator. Let us also mention that the temperature $\mathbf{T}$ of the plasma is defined as $\dfrac{3}{2}\boldsymbol{\rho}\mathbf{T} = \boldsymbol{\rho}\mathbf{E}$.

2. Any function of the type $\psi(\varepsilon) = \alpha\exp(-\beta\varepsilon)$ where α and β are arbitrary constants ($\beta > 0$) is a stationary solution of equation (2.1). The laws of conservation select, for this two-parameter set of functions, the unique equilibrium solution corresponding to the initial condition $f(\varepsilon,0) = f_0(\varepsilon)$.

3. The entropy defined by

$$H(f) \;=\; \int_0^\infty f(\varepsilon)\,\log(f(\varepsilon))\,\sqrt{\varepsilon}\,d\varepsilon\,, \tag{2.7}$$

   satisfies the classical H-Theorem. That means the entropy decays with time (by letting $\phi(\varepsilon) = \log(f(\varepsilon))$ in the weak formulation (2.4)) and

$$\frac{dH(f)}{dt} = 0 \iff f(\varepsilon) = \alpha \exp(-\beta\varepsilon)\,. \tag{2.8}$$

This is formally equivalent to say that f is an equilibrium function, that is $Q(f) = 0$.

The FPL operator can be rewritten in the following diffusive form

$$Q(f)(\varepsilon) = \frac{1}{\sqrt{\varepsilon}} \frac{\partial}{\partial\varepsilon} \left(E(f)\, f(\varepsilon) + D(f)\, \frac{\partial f}{\partial\varepsilon}\right), \tag{2.9}$$

where

$$\begin{aligned} D(f) &= \int_0^\infty g(\varepsilon, \varepsilon') f(\varepsilon') d\varepsilon'\,, \\ \text{and} & \\ E(f) &= -\int_0^\infty g(\varepsilon, \varepsilon') \frac{\partial f(\varepsilon')}{\partial \varepsilon'} d\varepsilon'\,. \end{aligned} \tag{2.10}$$

Remark 1. *To preserve the number of particles and the energy, Bobylev and Chuyanov [2] write the FPL operator in the form*

$$Q(f)(\varepsilon) = \frac{1}{\sqrt{\varepsilon}} \frac{\partial^2}{\partial\varepsilon^2} W(f, \varepsilon)\,,$$

where

$$W(f, \varepsilon) = \int_0^\varepsilon \int_0^\infty g(\varepsilon', \varepsilon'') \left(f(\varepsilon') \frac{\partial f(\varepsilon'')}{\partial \varepsilon''} - f(\varepsilon'') \frac{\partial f(\varepsilon')}{\partial \varepsilon'}\right) d\varepsilon'\, d\varepsilon''\,, \tag{2.11}$$

or, from (2.10)

$$W(f, \varepsilon) = \int_0^\varepsilon \left(E(f)\, f(\varepsilon') + D(f) \frac{\partial f(\varepsilon')}{\partial \varepsilon'}\right) d\varepsilon'\,. \tag{2.12}$$

By symmetry (between ε' and ε'' in (2.11)) $W(f, \varepsilon)$ vanishes at infinity and also at zero point, therefore density and energy are conserved.

Remark 2. *The mathematical analysis of the Landau equation is done by Desvillettes and Villani in [19, 20]. Concerning the derivation from the Boltzmann equation in the case of Coulombian potential there is the work of Degond and Lucquin [16]. There is also the work of Desvillettes for the derivation from the Boltzmann equation when the collisions become grazing [18]. For the development of the FPL operator in spherical harmonics we refer to the work of Shkarofsky et al. [30].*

2.1.1 Coulombian interactions

In the Coulombian case the collisions coefficients $E(f)$ and $D(f)$ describe drift and diffusion between electrons and are given by

$$E(f) = -\int_0^\varepsilon \varepsilon^{\frac{\prime 3}{2}} \frac{\partial f(\varepsilon')}{\partial \varepsilon'}\, d\varepsilon' - \varepsilon^{\frac{3}{2}} \int_\varepsilon^\infty \frac{\partial f(\varepsilon')}{\partial \varepsilon'}\, d\varepsilon' = -\int_0^\varepsilon \varepsilon^{\frac{\prime 3}{2}} \frac{\partial f(\varepsilon')}{\partial \varepsilon'}\, d\varepsilon' + \varepsilon^{\frac{3}{2}} f(\varepsilon)\,, \tag{2.13}$$

and

$$D(f) \,=\, \int_0^{\varepsilon} \varepsilon^{\frac{\imath 3}{2}}\, f(\varepsilon')\, d\varepsilon' \,+\, \varepsilon^{\frac{3}{2}} \int_{\varepsilon}^{\infty} f(\varepsilon')\, d\varepsilon'\,, \tag{2.14}$$

respectively. Integrating by parts and using the fact that f vanishes at ∞ we obtain

$$E(f) \,=\, \frac{3}{2} \int_0^{\varepsilon} \sqrt{\varepsilon'}\, f(\varepsilon')\, d\varepsilon'\,, \tag{2.15}$$

and

$$D(f) \,=\, \frac{3}{2} \int_0^{\varepsilon} \sqrt{\varepsilon'} \left(\int_{\varepsilon'}^{\infty} f(\varepsilon'')\, d\varepsilon'' \right) d\varepsilon'\,. \tag{2.16}$$

Deriving once $E(f)$ and twice $D(f)$ we have

$$f(\varepsilon) \,=\, \frac{2}{3}\, \frac{1}{\sqrt{\varepsilon}}\, \frac{\partial E(f)}{\partial \varepsilon}\,, \tag{2.17}$$

and

$$f(\varepsilon) \,=\, -\,\frac{2}{3}\, \frac{\partial}{\partial \varepsilon}\, \frac{1}{\sqrt{\varepsilon}}\, \frac{\partial}{\partial \varepsilon} D(f)\,. \tag{2.18}$$

Remark 3. *Since f is positive, $E(f)$ and $D(f)$ are positive.*

Remark 4. *At infinity, $E(f)$ is a density.*

Remark 5. *From relations (2.12), (2.15) and (2.16) we get in the Coulombian case*

$$W(f,\, \varepsilon) \,=\, \int_0^{\varepsilon} \left(\frac{3}{2} \int_0^{\varepsilon'} \sqrt{\varepsilon''}\, f(\varepsilon'') d\varepsilon'' \right) f(\varepsilon')\, d\varepsilon' \,+\, \int_0^{\varepsilon} D(f,\, \varepsilon')\, \frac{\partial f(\varepsilon')}{\partial \varepsilon'}\, d\varepsilon'\,.$$

By integrating by parts, we can write

$$\begin{aligned} W(f,\, \varepsilon) \,=& \frac{3}{2} \int_0^{\varepsilon} \sqrt{\varepsilon'} f(\varepsilon') \left(\int_{\varepsilon'}^{\infty} f(\varepsilon'') d\varepsilon'' \right) d\varepsilon' \,-\, \frac{3}{2} \left(\int_0^{\varepsilon} \sqrt{\varepsilon'} f(\varepsilon') d\varepsilon' \right) \left(\int_{\varepsilon}^{\infty} f(\varepsilon')\, d\varepsilon' \right) \\ & - \int_0^{\infty} \frac{\partial D(f, \varepsilon')}{\partial \varepsilon'} f(\varepsilon') d\varepsilon' \,+\, D(f,\, \varepsilon) f(\varepsilon)\,. \end{aligned}$$

As $\dfrac{\partial D(f,\varepsilon)}{\partial \varepsilon} = \dfrac{3}{2} \sqrt{\varepsilon} \displaystyle\int_{\varepsilon}^{\infty} f(\varepsilon') d\varepsilon'$ *the first and the third integral in the right-hand side of the previous equation become identical, thus we obtain*

$$W(f,\, \varepsilon) \,=\, -\,\frac{3}{2} \left(\int_0^{\varepsilon} \sqrt{\varepsilon'}\, f(\varepsilon')\, d\varepsilon' \right) \left(\int_{\varepsilon}^{\infty} f(\varepsilon')\, d\varepsilon' \right) \,+\, D(f,\, \varepsilon)\, f(\varepsilon)\,,$$

in other words

$$W(f,\, \varepsilon) \,=\, \frac{2}{3} \left(-\,\frac{E(f,\, \varepsilon)}{\sqrt{\varepsilon}}\, \frac{\partial}{\partial \varepsilon} D(f,\, \varepsilon) \,+\, \frac{D(f,\, \varepsilon)}{\sqrt{\varepsilon}}\, \frac{\partial}{\partial \varepsilon} E(f,\, \varepsilon) \right)\,.$$

2.1.2 Maxwellian interactions

In the Maxwellian interactions case we simply get

$$E(f) = \frac{3}{2}\varepsilon^{\frac{3}{2}}\boldsymbol{\rho} \text{ and } D(f) = \varepsilon^{\frac{3}{2}}\boldsymbol{\rho}\mathbf{E} = \frac{3}{2}\varepsilon^{\frac{3}{2}}\boldsymbol{\rho}\,\mathbf{T}\,.$$

Then, equation (2.9) is linear (in the sense that $E(f)$ and $D(f)$ are independent of f). We can write

$$Q(f)(\varepsilon) = \frac{3}{2}\frac{\boldsymbol{\rho}}{\sqrt{\varepsilon}}\frac{\partial}{\partial\varepsilon}\left(\varepsilon^{\frac{3}{2}}\,(\,f(\varepsilon) + \mathbf{T}\,\frac{\partial f(\varepsilon)}{\partial\varepsilon})\right).$$

3 Semi-discretized problem

In this section we focus on the discretization in the energy variable. For numerical simulations, we reduce the integration domain in FPL equation to a bounded domain in the variable ε where $\varepsilon \in [0, \mathcal{E}]$. Thus we consider the approximate problem of (2.3) defined by

$$\int_0^{\mathcal{E}} \frac{\partial f(\varepsilon,t)}{\partial t}\,\phi(\varepsilon)\sqrt{\varepsilon}d\varepsilon = -\frac{1}{2}\int_0^{\mathcal{E}}\int_0^{\mathcal{E}} (\frac{\partial\phi(\varepsilon)}{\partial\varepsilon} - \frac{\partial\phi(\varepsilon')}{\partial\varepsilon'})\,g(\varepsilon,\varepsilon')\,(f(\varepsilon')\frac{\partial f(\varepsilon)}{\partial\varepsilon} - f(\varepsilon)\frac{\partial f(\varepsilon')}{\partial\varepsilon'})\,d\varepsilon'\,d\varepsilon\,. \tag{3.1}$$

Let us introduce $\{\varepsilon_i\}_{1\le i\le N+1}$ an increasing sequence such that $\varepsilon_1 = 0$, $\varepsilon_N = \mathcal{E}$ and $\Delta\varepsilon_{i+\frac{1}{2}} = \varepsilon_{i+1} - \varepsilon_i$. We suppose that $\{\Delta\varepsilon_{i+\frac{1}{2}}\}_{1\le i\le N}$ is constant or increasing sequence. We define $v_{i+\frac{1}{2}}$ as the mean value of the velocity on $[\varepsilon_i, \varepsilon_{i+1}]$ i.e.

$$v_{i+\frac{1}{2}} = \frac{1}{\Delta\varepsilon_{i+\frac{1}{2}}}\int_{\varepsilon_i}^{\varepsilon_{i+1}}\sqrt{\varepsilon}\,d\varepsilon = \frac{2}{3\Delta\varepsilon_{i+\frac{1}{2}}}\,(\varepsilon_{i+1}^{\frac{3}{2}} - \varepsilon_i^{\frac{3}{2}})\,,$$

so, we get $\varepsilon_{i+\frac{1}{2}}$ and can define

$$(v^2\Delta v)_i = \frac{1}{2}\int_{\varepsilon_{i-\frac{1}{2}}}^{\varepsilon_{i+\frac{1}{2}}}\sqrt{\varepsilon}d\varepsilon = \frac{1}{2}\sqrt{\varepsilon_i}\,\Delta\varepsilon_i = \frac{1}{3}(\varepsilon_{i+\frac{1}{2}}^{\frac{3}{2}} - \varepsilon_{i-\frac{1}{2}}^{\frac{3}{2}})\,. \tag{3.2}$$

By convention we set $\varepsilon_{\frac{1}{2}} = 0$. Any function $f(\varepsilon,t)$ is approximated on the grid by values $\{f_i\}_{1\le i\le N}$ supposed be approximations of $\{f(\varepsilon_i)\}_{1\le i\le N}$. We also introduce the notations $(\Delta\phi)_{i+\frac{1}{2}} = \phi_{i+1} - \phi_i$ and $(\mathbb{D}\phi)_{i+\frac{1}{2}} = \dfrac{(\Delta\phi)_{i+\frac{1}{2}}}{(\Delta\varepsilon)_{i+\frac{1}{2}}}$ as an approximation of the partial derivative $\partial_\varepsilon\phi(\varepsilon_{i+\frac{1}{2}})$.

First, we consider the discretization of the expression $\displaystyle\int_0^{\mathcal{E}}\frac{\partial f}{\partial t}\phi(\varepsilon)\sqrt{\varepsilon}d\varepsilon$ for any function ϕ. By writing

$$\int_0^{\mathcal{E}}\frac{\partial f}{\partial t}\phi\sqrt{\varepsilon}d\varepsilon \simeq \sum_{i=1}^{N}\int_{\varepsilon_{i-\frac{1}{2}}}^{\varepsilon_{i+\frac{1}{2}}}\frac{\partial f}{\partial t}\phi(\varepsilon)\,\sqrt{\varepsilon}\,d\varepsilon,$$

and using the standard quadrature formula with respect to the measure $\sqrt{\varepsilon}d\varepsilon$, we approximate it by

$$\sum_{i=1}^{N}(\phi_i \frac{df_i}{dt} \int_{\varepsilon_{i-\frac{1}{2}}}^{\varepsilon_{i+\frac{1}{2}}} \sqrt{\varepsilon}\, d\varepsilon) \stackrel{def}{=} \sum_{i=1}^{N} c_i \phi_i \frac{df_i}{dt} \tag{3.3}$$

Thus the weights c_i are defined by $c_1 = \frac{2}{3}\varepsilon_{\frac{3}{2}}^{\frac{3}{2}}$, $c_i = \frac{2}{3}(\varepsilon_{i+\frac{1}{2}}^{\frac{3}{2}} - \varepsilon_{i-\frac{1}{2}}^{\frac{3}{2}})$ for $i = 2, \cdots, N$.
Below, we present various strategies to construct schemes that have properties of conservation, and if possible positivity and entropy decaying. These schemes differ intrinsically in the way we discretize the right-hand side of (3.1):

$$(r.h.s.) \;=\; -\frac{1}{2}\sum_{i=1}^{N-1}\sum_{j=1}^{N-1}\left(\int_{\varepsilon_i}^{\varepsilon_{i+1}}\int_{\varepsilon_j}^{\varepsilon_{j+1}} (\frac{\partial\phi(\varepsilon)}{\partial\varepsilon} - \frac{\partial\phi(\varepsilon')}{\partial\varepsilon'})\, g(\varepsilon,\varepsilon')\, (f(\varepsilon')\frac{\partial f(\varepsilon)}{\partial\varepsilon} \;-\; f(\varepsilon)\frac{\partial f(\varepsilon')}{\partial\varepsilon'})\, d\varepsilon\, d\varepsilon'\right). \tag{3.4}$$

We must also introduce the Maxwellian associated to a distribution function f. For a distribution function f we define $\overline{\boldsymbol{\rho}}$ and $\overline{\boldsymbol{\rho}\mathbf{E}}$ the discretized analogous of density (2.5) and energy (2.6) as

$$\overline{\boldsymbol{\rho}} \;=\; \sum_{i=1}^{N} c_i\, f_i \;, \quad \overline{\boldsymbol{\rho}\mathbf{E}} \;=\; \sum_{i=1}^{N} \varepsilon_i\, c_i\, f_i \,.$$

The temperature is still defined by $\frac{3}{2}\overline{\boldsymbol{\rho}}\,\overline{\mathbf{T}} = \overline{\boldsymbol{\rho}\mathbf{E}}$. For a distribution function f, we denote by $M = \alpha\exp(-\beta\varepsilon)$ the Maxwellian which has the same mass and energy as f. It's easy to check that

$$\widetilde{T} = \frac{\sum_{i=1}^{N}\ \varepsilon_i\ c_i\, M_i}{\sum_{i=1}^{N}\ c_i\, M_i} = \frac{\sum_{i=1}^{N}\ \varepsilon_i\ c_i\ \exp(-\beta\varepsilon_i)}{\sum_{i=1}^{N}\ c_i\ \exp(-\beta\varepsilon_i)}$$

is a strictly monotone (decreasing) function of β with

$$\lim_{\beta\to 0}\widetilde{T} = T_{\max} = \frac{\sum_{i=1}^{N}\ \varepsilon_i\ c_i}{\sum_{i=1}^{N}\ c_i}$$

and

$$\lim_{\beta\to+\infty}\widetilde{T} = T_{\min} = 0.$$

Thus for any distribution function f such that $0 \le \overline{\mathbf{T}} \le T_{\max}$ there is a unique $\beta \ge 0$ such that $\widetilde{T} = \overline{\mathbf{T}}$ and consequently M is unique. Note that $\beta \simeq \dfrac{1}{\overline{\mathbf{T}}}$. For the rest of this work we consider only distribution functions such that $0 \le \overline{\mathbf{T}} \le T_{\max}$, that is, we exclude distribution function for which M is an increasing function of ε ($\beta \le 0$)(that means we exclude distribution function such that $T_{\max} \le \overline{\mathbf{T}} \le \dfrac{2}{3}\mathcal{E}$).

To check the positivity of the schemes we will have the need of the following well-known result:

Lemma 1. *Consider the Cauchy problem for the ordinary differential equation*

$$\frac{df}{dt} = Lf$$

$$f(t=0) = f^0$$

with $f = \{f_i\}_{1 \leq i \leq N}$ *and with the square matrix* $L = L(f,t)$ *such that* $L_{ij} \geq 0$ *for* $i \neq j$, $L_{ii} \leq 0$.

If there exists a constant C *such that* $\forall\ i,j\ |L_{ij}(f,t)| \leq C$ *then, if the solution exists on a time interval* $[0,T]$ *and if the initial data is positive i.e.* $f^0 > 0$, *the solution* f *is positive i.e.* $f_i(t) > 0$ *for all* i *and for all time* $t \in [0,T]$.

If there exists a constant C *such that* $\forall\ i,j\ |L_{ij}(f,t)| \leq C$ *then, if the solution exists on a time interval* $[0,T]$ *and if the initial data is non-negative i.e.* $f^0 \geq 0$, *the solution* f *is non-negative i.e.* $f_i(t) \geq 0$ *for all* i *and for all time* $t \in [0,T]$.

Such an assumption for the matrix L ensures the existence and the uniqueness of a global solution in time if the matrix L is locally Lipchitz in f. If the matrix L is only continuous in f we have also the existence of a global solution in time but not the uniqueness. One can see [13] for these well known results. Let us also recall that the solution in these cases is C^1 in the time variable.

Proof. First we suppose that $f^0 > 0$. Let $t_0 \leq \tau$ be the first time for which there exists an index i_0 such that $f_{i_0}(t_0) = 0$. In $[0,t_0]$ for all i, we have $\dfrac{df_i}{dt} \geq -Cf_i$ and by integrating $f_i(t) \geq f_i^0 \exp(-Ct)$. Thus the solution cannot vanish in finite time. The extension of this result to the case of a non-negative initial data ($f_i^0 \geq 0$ for all i) depends on the regularity of the matrix L with respect to f.

If L is Lipschitz in f the solution depends continuously on the initial data, see by example [13], and thus by taking a sequence $f_\nu^0 > 0$ such that $f_\nu^0 \to f^0$ and by using the result above concerning strictly positive initial data, f is non-negative.

We suppose now that L is just continuous in f. First we assume that for all f and for all i, $|L_{ii}| > \sum_{j \neq i} L_{ij}$ that is $L(f)$ is strictly diagonally dominant. We suppose that the solution becomes negative for a set I of indexes up to a time $t_1 > 0$ ($I \subset \{i \text{ such that } \ f_i^0 = 0\}$). Since the solution is continuous in time, $\min_i f_i$ is defined and continuous in time in $[0,t_1]$. Let $J \subset I$ such that $f_j = \min_i f_i$ for all $j \in J$ and for all $t \in [0,t_2]$, $t_2 \leq t_1$.
We have the following lowerbound for $j \in J$:

$$f_j(t_1) = \int_0^{t_1} (\sum_{k \neq j} L_{jk}(f) f_k + L_{jj}(f) f_j) dt \geq \int_0^{t_1} (\sum_{k \neq j, k \in I} L_{jk}(f) f_k + L_{jj}(f) f_j) dt > 0$$

thus this is in contradiction with our assumption that in $[0,t_1]$, $f_j < 0$. Thus the solution verifies $f \geq 0$.
Now if $L(f)$ is not strictly diagonally dominant but verifies $|L_{ij}(f,t)| \leq C$, we can choose a constant K such that $L(f) - K\mathbf{Id}$ is strictly diagonally dominant. We have

$$\frac{df}{dt} = Lf \iff \frac{df}{dt} - Kf = (L - K\mathbf{Id})f,$$

where **Id** is the identity matrix. Thus if we set $g = \exp(-Kt)f$, g verifies

$$\frac{df}{dt} = (L(g\exp(Kt)) - K\mathbf{Id})g = \mathcal{L}(g)g,$$

and the matrix $\mathcal{L}(g)$ is strictly diagonally dominant. Using the above result for strictly diagonally dominant matrix we have $g \geq 0$ and thus $f \geq 0$.
If we do not suppose that $L(f)$ is continuous in f, but if we suppose the existence of a solution continuous in time in a time interval $[0, T]$, the proof above is still valid on this time interval. Thus the solution is non-negative on $[0, T]$.
This ends the proof. □

3.1 Chang and Cooper type schemes

One of the most popular method used for Fokker-Planck equations is due to Chang and Cooper [10]. Originally this method was proposed for the linear Fokker-Planck equation and the construction is entirely devoted to the preservation of equilibrium states. The authors also show that in the linear case (and only in this case) the method provides also non-negative solutions. At contrary others authors, see [25], have shown that for the non-linear Fokker-Planck equation this method can produce non-positive solutions. Let us recall the spirit of the Chang and Cooper method on the simple Fokker-Planck equation for a distribution function $f(v,t)$, $v \in \mathbb{R}$,

$$\frac{\partial f}{\partial t} = \frac{\partial F}{\partial v} = \frac{\partial}{\partial v}(vf + \sigma\frac{\partial f}{\partial v}),$$

that can be put under the form

$$\frac{\partial f}{\partial t} = \sigma\frac{\partial}{\partial v}(M\frac{\partial}{\partial v}(\frac{f}{M})),$$

where M is a Maxwellian, and $M = M(v) = \exp(-|v|^2/2\sigma)$. Note that M is the equilibrium, the long time behavior of the distribution function. When $f = M$, F vanishes.
On a uniform grid of velocity space step Δv and $v_i = i\Delta v$, the Chang and Cooper method consists to discretize the diffusion as usual and the drift in such a way that for $f = M$ the fluxes F are all equal to zero. One takes at cell interface $v_{i+\frac{1}{2}} = (i + \frac{1}{2})\Delta v$

$$F_{i+\frac{1}{2}} = \frac{\sigma}{\Delta v}(f_{i+1} - f_i) + v_{i+\frac{1}{2}}\left((1 - \delta_{i+\frac{1}{2}})f_{i+1} + \delta_{i+\frac{1}{2}}f_i\right),$$

and forcing the fluxes to be zero for equilibrium leads to the following definition of $\delta_{i+\frac{1}{2}}$

$$\delta_{i+\frac{1}{2}} = \frac{\sigma}{v_{i+\frac{1}{2}}\Delta v} - \frac{1}{\exp(\frac{v_{i+\frac{1}{2}}\Delta v}{\sigma}) - 1}.$$

The scheme writes as

$$\frac{df_i}{dt} = \frac{1}{\Delta v}(F_{i+\frac{1}{2}} - F_{i-\frac{1}{2}}).$$

As explained in [8, 9] this scheme is in fact an entropic scheme for the above example.

This method is also one of the most used for FPL equation. And this is a non-linear problem. None of the main work in this area [11, 22, 23, 24] contains the proof of the positivity of the Chang and Cooper method or the proof of the energy conservation. Moreover it is not clear that equilibrium states are preserved by this method for the FPL equation since the definition of the coefficients of Chang and Cooper, see [10] of these coefficients, are defined in an implicit manner.
In this section we propose a new derivation of this method which intrinsically contains the conservation of the energy. We recall also the scheme developed by Langdon *et al.* [15, 24], but in the case where potentials are Coulombian, and extend it to the general case. We show clearly why our new method provides non-negative solution and why equilibrium states are still preserved. As we see later, this analysis clarifies the confusion about the boundary condition used at one end of the domain of computation, more precisely in $\varepsilon = \mathcal{E}$. In particular this analysis shows that, due to boundary condition at $\varepsilon = \mathcal{E}$, we can have some doubt about the energy conservation for the version of the scheme used by Epperlein in [22] and also used by other people, [23] for example. The discretization is only in energy, but the analysis provides also some results about implicit time discretization, more precisely about some simplifications of the totally implicit time discretization which could be made: the only positive, conservative and equilibrium state preserving time implicit discretization of the Chang and Cooper method for the FPL equation is the fully implicit one.
As example, the schemes are applied to the case where potentials are Coulombian or Maxwellian.

3.1.1 A new variant: scheme $\mathcal{S}_1$

Using for each integrals of (3.4) a midpoint quadrature formula, we approximate it by

$$(r.h.s.) \;=\; -\frac{1}{2}\sum_{i=1}^{N-1}\sum_{j=1}^{N-1}((\mathbb{D}\phi)_{i+\frac{1}{2}} - (\mathbb{D}\phi)_{j+\frac{1}{2}})\, g_{i+\frac{1}{2},j+\frac{1}{2}}(f_{j+\frac{1}{2}}(\mathbb{D}f)_{i+\frac{1}{2}} - f_{i+\frac{1}{2}}(\mathbb{D}f)_{j+\frac{1}{2}})\,\Delta\varepsilon_{i+\frac{1}{2}}\,\Delta\varepsilon_{j+\frac{1}{2}}\,, \tag{3.5}$$

or

$$(r.h.s.) \;=\; -\sum_{i=1}^{N-1}(\mathbb{D}\phi)_{i+\frac{1}{2}}\,\Delta\varepsilon_{i+\frac{1}{2}}\sum_{j=1}^{N-1} g_{i+\frac{1}{2},j+\frac{1}{2}}\,(f_{j+\frac{1}{2}}\,(\mathbb{D}f)_{i+\frac{1}{2}} \;-\; f_{i+\frac{1}{2}}\,(\mathbb{D}f)_{j+\frac{1}{2}})\,\Delta\varepsilon_{j+\frac{1}{2}}\,,$$

with $g_{i+\frac{1}{2},j+\frac{1}{2}} = g(\varepsilon_{i+\frac{1}{2}}, \varepsilon_{j+\frac{1}{2}})$. Hence, the weak formulation of the semi-discretized model reads

$$\sum_{i=1}^{N} c_i \frac{\partial f_i}{\partial t}\phi_i \;=\; -\sum_{i=1}^{N-1}(\mathbb{D}\phi)_{i+\frac{1}{2}}\,\Delta\varepsilon_{i+\frac{1}{2}}\sum_{j=1}^{N-1} g_{i+\frac{1}{2},j+\frac{1}{2}}\,(f_{j+\frac{1}{2}}\,(\mathbb{D}f)_{i+\frac{1}{2}} \;-\; f_{i+\frac{1}{2}}\,(\mathbb{D}f)_{j+\frac{1}{2}})\,\Delta\varepsilon_{j+\frac{1}{2}}\,. \tag{3.6}$$

To simplify we note $K_{i+\frac{1}{2}} = \displaystyle\sum_{j=1}^{N-1} g_{i+\frac{1}{2},j+\frac{1}{2}}\,(f_{j+\frac{1}{2}}\,(\mathbb{D}f)_{i+\frac{1}{2}} \;-\; f_{i+\frac{1}{2}}\,(\mathbb{D}f)_{j+\frac{1}{2}})\,\Delta\varepsilon_{j+\frac{1}{2}}$ the numerical flux. By factorizing the term ϕ_i in the right-hand side of (3.6), we have

$$\sum_{i=1}^{N} c_i \frac{\partial f_i}{\partial t} \phi_i \;=\; \sum_{j=2}^{N-1} (K_{j+\frac{1}{2}} \;-\; K_{j-\frac{1}{2}})\, \phi_j \;+\; \phi_1\, K_{\frac{3}{2}} \;-\; \phi_N\, K_{N-\frac{1}{2}}\,.$$

Finally, by identifying the terms involving ϕ_i in (3.6), we obtain the system of ordinary differential equation

$$\frac{df_i}{dt} \;=\; Q_i^{\mathcal{S}_1} \qquad 1 \;\leq\; i \;\leq N \tag{3.7}$$

with $Q_1^{\mathcal{S}_1} = K_{\frac{3}{2}}/c_1$, $Q_i^{\mathcal{S}_1} = (K_{i+\frac{1}{2}} - K_{i-\frac{1}{2}})/c_i$ for $2 \leq i \leq N-1$ and $Q_N^{\mathcal{S}_1} = -K_{N-\frac{1}{2}}/c_N$. We can rewrite the numerical flux

$$K_{i+\frac{1}{2}} \;=\; -\, f_{i+\frac{1}{2}} \sum_{j=1}^{N-1} g_{i+\frac{1}{2},j+\frac{1}{2}}\, (\mathbb{D}f)_{j+\frac{1}{2}} \Delta\varepsilon_{j+\frac{1}{2}} + (\mathbb{D}f)_{i+\frac{1}{2}} \sum_{j=1}^{N-1} g_{i+\frac{1}{2},j+\frac{1}{2}}\, f_{j+\frac{1}{2}}\, \Delta\varepsilon_{j+\frac{1}{2}}\,, \quad 1 \;\leq\; i \;\leq\; N-1 \tag{3.8}$$

and recognize the discretized collisions terms (2.10)

$$E_{i+\frac{1}{2}} \;=\; -\sum_{j=1}^{N-1} g_{i+\frac{1}{2},j+\frac{1}{2}}\, (\mathbb{D}f)_{j+\frac{1}{2}}\, \Delta\varepsilon_{j+\frac{1}{2}} \quad \text{and} \quad D_{i+\frac{1}{2}} \;=\; \sum_{j=1}^{N-1} g_{i+\frac{1}{2},j+\frac{1}{2}}\, f_{j+\frac{1}{2}}\, \Delta\varepsilon_{j+\frac{1}{2}}\,, \tag{3.9}$$

therefore

$$K_{i+\frac{1}{2}} \;=\; E_{i+\frac{1}{2}}\, f_{i+\frac{1}{2}} \;+\; D_{i+\frac{1}{2}}\, \frac{f_{i+1} \;-\; f_i}{\Delta\varepsilon_{i+\frac{1}{2}}}\,, \quad 1 \leq i \leq N-1 \tag{3.10}$$

Remark 6. *By integrating by parts the drift term reads*

$$E_{i+\frac{1}{2}} \;=\; \sum_{j=1}^{N-1} (g_{i+\frac{1}{2},j+\frac{1}{2}} \;-\; g_{i+\frac{1}{2},j-\frac{1}{2}})\, f_j \;-\; g_{i+\frac{1}{2},N-\frac{1}{2}}\, f_N\,.$$

We assume that $\{g_{i+\frac{1}{2},j+\frac{1}{2}}\}_{1\leq i\leq N}$ is an increasing sequence. Then, even if the f_i's are positive, $E_{i+\frac{1}{2}}$ can be negative if $f_N \neq 0$. In other hand, since the $f_{i+\frac{1}{2}}$'s are positive, $D_{i+\frac{1}{2}}$ remains positive.

We turn now to the discretization of $f_{i+\frac{1}{2}}$. We suppose that $f_{i+\frac{1}{2}}$ is an approximation of $f(\varepsilon_{i+\frac{1}{2}})$ given by the following definition [10]

Definition 1. *The Chang and Cooper average $f_{i+\frac{1}{2}}$ of quantities f_i and f_{i+1} is defined by*

$$f_{i+\frac{1}{2}} \;=\; \delta_{i+\frac{1}{2}}\, f_i \;+\; (1 \;-\; \delta_{i+\frac{1}{2}})\, f_{i+1} \qquad 1 \leq i \leq N-1\,, \tag{3.11}$$

with

$$\delta_{i+\frac{1}{2}} = \frac{1}{\alpha_{i+\frac{1}{2}}} - \frac{1}{\exp(\alpha_{i+\frac{1}{2}}) - 1} \qquad 1 \leq i \leq N-1\,, \tag{3.12}$$

and $\alpha_{i+\frac{1}{2}} = \dfrac{E_{i+\frac{1}{2}}}{D_{i+\frac{1}{2}}} \Delta\varepsilon_{i+\frac{1}{2}}$.

To be comfortable we denote $h(\alpha) = \dfrac{1}{\alpha} - \dfrac{1}{(\exp(\alpha)-1)}$ (thus $\delta_{i+\frac{1}{2}} = h(\alpha_{i+\frac{1}{2}})$). This smooth function is decreasing and bounded such that $h(-\infty) = 1$, $h(0) = 1/2$ and $h(+\infty) = 0$. Moreover, his derivative is negative and bounded too ($h'(\pm\infty) = 0$). Note that whatever the value of $\alpha_{i+\frac{1}{2}}$ (positive or negative) if $\delta_{i+\frac{1}{2}}$ exists we get $0 \leq \delta_{i+\frac{1}{2}} \leq 1$. Therefore, since the f_i's are positive the $f_{i+\frac{1}{2}}$'s remain positive. Now we clarify the expression of $\delta_{i+\frac{1}{2}}$. We recall that f is the N-dimensional column vector with components $\{f_i\}_{1\leq i\leq N}$. First $\alpha_{i+\frac{1}{2}}$ (for $1 \leq i \leq N-1$) reads

$$\alpha_{i+\frac{1}{2}} \;=\; \frac{E_{i+\frac{1}{2}}}{D_{i+\frac{1}{2}}}\Delta\varepsilon_{i+\frac{1}{2}} \;=\; \left(\frac{-\displaystyle\sum_{j=1}^{N-1} g_{i+\frac{1}{2},j+\frac{1}{2}}\,(\mathbb{D}f)_{j+\frac{1}{2}}\,\Delta\varepsilon_{j+\frac{1}{2}}}{\displaystyle\sum_{j=1}^{N-1} g_{i+\frac{1}{2},j+\frac{1}{2}}\,f_{j+\frac{1}{2}}\,\Delta\varepsilon_{j+\frac{1}{2}}} \right) \Delta\varepsilon_{i+\frac{1}{2}}\,,$$

$$\alpha_{i+\frac{1}{2}} \;=\; \left(\frac{-\displaystyle\sum_{j=1}^{N-1} g_{i+\frac{1}{2},j+\frac{1}{2}}\,(\mathbb{D}f)_{j+\frac{1}{2}}\,\Delta\varepsilon_{j+\frac{1}{2}}}{\displaystyle\sum_{j=1}^{N-1} g_{i+\frac{1}{2},j+\frac{1}{2}}\,((1-\delta_{j+\frac{1}{2}})f_{j+1} + \delta_{j+\frac{1}{2}}f_j)\,\Delta\varepsilon_{j+\frac{1}{2}}} \right) \Delta\varepsilon_{i+\frac{1}{2}}\,,$$

and $\delta_{i+\frac{1}{2}} = h(\alpha_{i+\frac{1}{2}})$. Thus, assuming that δ is the (N-1)-dimensional column vector with components $\{\delta_{i+\frac{1}{2}}\}_{1\leq i\leq N-1}$, we can write $\delta_{i+\frac{1}{2}}$ as a function of f and introducing $\{\mathcal{H}_{i+\frac{1}{2}}\}_{1\leq i\leq N-1}$ we get

$$\delta_{i+\frac{1}{2}}(f) \;=\; \mathcal{H}_{i+\frac{1}{2}}(\delta_{\frac{3}{2}},\, ...,\, \delta_{k+\frac{1}{2}},\, ...,\, \delta_{N-\frac{1}{2}},\, f_1,\, ...,\, f_k,\, ...\, f_N)\,,$$

in other words, the system to be solved is

$$\delta(f) = \mathcal{H}(\delta(f), f)\,, \tag{3.13}$$

where f is a distribution function that has density $\overline{\boldsymbol{\rho}}$ and energy $\overline{\boldsymbol{\rho}\mathbf{E}}$ and $\mathcal{H}$ is a (N-1)-dimensional vector valued function. Before solving this non-linear problem, we have to prove the existence of solution. To begin with, as f is a distribution function that has density $\overline{\boldsymbol{\rho}}$ and energy $\overline{\boldsymbol{\rho}\mathbf{E}}$, $\{E_{i+\frac{1}{2}}\}_{1\leq i\leq N-1}$ and $\{D_{i+\frac{1}{2}}\}_{1\leq i\leq N-1}$ are well defined and continuous in respect to δ. Hence $\{\alpha_{i+\frac{1}{2}}\}_{1\leq i\leq N-1}$ exists (positive or negative) and consequently $\mathcal{H}$ lies in $[0,1]^{N-1}$. Because h is continuous, $\mathcal{H}$ is a continuous mapping which carries $[0,1]^{N-1}$ into itself and thanks to Brouwer's theorem $\mathcal{H}$ has a fixed point (i.e. there exists an $\overline{\delta}(f)$ with $\mathcal{H}(\overline{\delta}(f), f) = \overline{\delta}(f)$). Unfortunately we don't get actually an effective proof for the uniqueness of solution (as, for example, the Picard's theorem). Nevertheless, we assume that we have sufficient conditions for the uniqueness of solution of the non-linear operator and that under these conditions the sequence generated by Newton's method converges to the solution. In each step of Newton's process we have to solve a linear system which associated matrix is non-sparse. That requires $\mathcal{O}(N^2)$ operations. But we will see that for Maxwellian and Coulombian potential this cost can be reduce to $\mathcal{O}(N)$ operations.

Proposition 1. *The flux $K_{i+\frac{1}{2}}$ satisfies the following relation*

$$K_{i+\frac{1}{2}} \,=\, A_{i+\frac{1}{2}}\, f_{i+1} \,-\, B_{i+\frac{1}{2}}\, f_i \quad , \quad \forall i\,;\ 1 \le i \le N-1 \tag{3.14}$$

where

$$A_{i+\frac{1}{2}} = \, v(\alpha_{i+\frac{1}{2}})\, \frac{D_{i+\frac{1}{2}}}{\Delta\varepsilon_{i+\frac{1}{2}}}\,, \tag{3.15}$$

and

$$B_{i+\frac{1}{2}} = \, u(\alpha_{i+\frac{1}{2}})\, \frac{D_{i+\frac{1}{2}}}{\Delta\varepsilon_{i+\frac{1}{2}}}\,, \tag{3.16}$$

with $u(\alpha) = \dfrac{\alpha}{\exp(\alpha)-1}$ *and* $v(\alpha) = \dfrac{\alpha\exp(\alpha)}{\exp(\alpha)-1}$.

Proof. The proof is the same as in Chang and Cooper's paper [10] but in the non-linear case. By substituting the Chang and Cooper average $f_{i+\frac{1}{2}}$ in the right-hand side of the relation (3.10) we get

$$K_{i+\frac{1}{2}} \,=\, E_{i+\frac{1}{2}}\,((1\,-\,\delta_{i+\frac{1}{2}})\, f_{i+1} \,+\, \delta_{i+\frac{1}{2}}\, f_i) \,+\, D_{i+\frac{1}{2}}\, \frac{f_{i+1}\,-\,f_i}{\Delta\varepsilon_{i+\frac{1}{2}}} \qquad \forall i\,;\ 1 \le i \le N-1\,.$$

As $\alpha_{i+\frac{1}{2}} = \dfrac{E_{i+\frac{1}{2}}}{D_{i+\frac{1}{2}}}\Delta\varepsilon_{i+\frac{1}{2}}$, we can write $E_{i+\frac{1}{2}} = \dfrac{\alpha_{i+\frac{1}{2}} D_{i+\frac{1}{2}}}{\Delta\varepsilon_{i+\frac{1}{2}}}$. Thus

$$K_{i+\frac{1}{2}} \,=\, \alpha_{i+\frac{1}{2}}\, \frac{D_{i+\frac{1}{2}}}{\Delta\varepsilon_{i+\frac{1}{2}}}((1-\delta_{i+\frac{1}{2}})f_{i+1} \,+\, \delta_{i+\frac{1}{2}} f_i) \,+\, D_{i+\frac{1}{2}}\, \frac{f_{i+1}\,-\,f_i}{\Delta\varepsilon_{i+\frac{1}{2}}}\,,$$

that means

$$K_{i+\frac{1}{2}} \,=\, (\alpha_{i+\frac{1}{2}}(1-\delta_{i+\frac{1}{2}}) \,+\, 1)\, \frac{D_{i+\frac{1}{2}}}{\Delta\varepsilon_{i+\frac{1}{2}}}\, f_{i+1} \,+\, (\alpha_{i+\frac{1}{2}}\delta_{i+\frac{1}{2}} - 1)\frac{D_{i+\frac{1}{2}}}{\Delta\varepsilon_{i+\frac{1}{2}}}\, f_i\,.$$

Finally, if we develop $\delta_{i+\frac{1}{2}}$ we have

$$K_{i+\frac{1}{2}} \,=\, \frac{\alpha_{i+\frac{1}{2}}\exp(\alpha_{i+\frac{1}{2}})}{\exp(\alpha_{i+\frac{1}{2}})-1}\, \frac{D_{i+\frac{1}{2}}}{\Delta\varepsilon_{i+\frac{1}{2}}}\, f_{i+1} \,-\, \frac{\alpha_{i+\frac{1}{2}}}{\exp(\alpha_{i+\frac{1}{2}})-1}\, \frac{D_{i+\frac{1}{2}}}{\Delta\varepsilon_{i+\frac{1}{2}}}\, f_i\,.$$

And now, we can write $K_{i+\frac{1}{2}} = A_{i+\frac{1}{2}} f_{i+1} - B_{i+\frac{1}{2}} f_i$. □

Summary To summarize we name scheme $\mathcal{S}_1$ the discretized system (3.7) where the numerical fluxes are (for $1 \le i \le N-1$)

$K_{i+\frac{1}{2}} = E_{i+\frac{1}{2}} f_{i+\frac{1}{2}} + D_{i+\frac{1}{2}}(f_{i+1}-f_i)/\Delta\varepsilon_{i+\frac{1}{2}}$ with $E_{i+\frac{1}{2}} = -\displaystyle\sum_{j=1}^{N-1} g_{i+\frac{1}{2},j+\frac{1}{2}}(\mathbb{D}f)_{j+\frac{1}{2}}\Delta\varepsilon_{j+\frac{1}{2}}$,

$D_{i+\frac{1}{2}} = \displaystyle\sum_{j=1}^{N-1} g_{i+\frac{1}{2},j+\frac{1}{2}} f_{j+\frac{1}{2}}\Delta\varepsilon_{j+\frac{1}{2}}$ and $f_{i+\frac{1}{2}}$ is given by the Chang and Cooper average (3.11)-(3.12).

3.1.2 The scheme $\mathcal{S}_2$ (Langdon's scheme in the case of a Coulombian potential)

Another option consists in integrating the right-hand side of (3.6) up to ε_{N+1}, assuming that $f_{N+1} = 0$ and constraining the flux at the last point $\varepsilon_{N+\frac{1}{2}}$ to be identically equal to zero. This strategy was carried by Langdon [24] and Decoster and Langdon [15] in the case of Coulombian interactions. We extend their method to the general case. We obtain

$$(r.h.s.) \;=\; -\,\frac{1}{2}\,\sum_{i=1}^{N}\sum_{j=1}^{N}((\mathbb{D}\phi)_{i+\frac{1}{2}} - (\mathbb{D}\phi)_{j+\frac{1}{2}})\,g_{i+\frac{1}{2},j+\frac{1}{2}}(f_{j+\frac{1}{2}}(\mathbb{D}f)_{i+\frac{1}{2}} - f_{i+\frac{1}{2}}(\mathbb{D}f)_{j+\frac{1}{2}})\,\Delta\varepsilon_{i+\frac{1}{2}}\,\Delta\varepsilon_{j+\frac{1}{2}}\,, \tag{3.17}$$

and

$$\sum_{i=1}^{N} c_i\frac{\partial f_i}{\partial t}\phi_i \;=\; -\sum_{i=1}^{N}(\mathbb{D}\phi)_{i+\frac{1}{2}}\,\Delta\varepsilon_{i+\frac{1}{2}}\,\sum_{j=1}^{N} g_{i+\frac{1}{2},j+\frac{1}{2}}\,(f_{j+\frac{1}{2}}\,(\mathbb{D}f)_{i+\frac{1}{2}} \;-\; f_{i+\frac{1}{2}}\,(\mathbb{D}f)_{j+\frac{1}{2}})\,\Delta\varepsilon_{j+\frac{1}{2}}\,. \tag{3.18}$$

Finally, we get the weak formulation of the semi-discretized model

$$\frac{df_i}{dt} \;=\; Q_i^{\mathcal{S}_2} \qquad 1 \;\le\; i \;\le N \tag{3.19}$$

with $Q_1^{\mathcal{S}_2} = \widetilde{K}_{\frac{3}{2}}/c_1$, $Q_i^{\mathcal{S}_2} = (\widetilde{K}_{i+\frac{1}{2}} - \widetilde{K}_{i-\frac{1}{2}})/c_i$ for $2 \le i \le N$, where

$$\widetilde{K}_{i+\frac{1}{2}} \;=\; -\,f_{i+\frac{1}{2}}\,\sum_{j=1}^{N} g_{i+\frac{1}{2},j+\frac{1}{2}}\,(\mathbb{D}f)_{j+\frac{1}{2}}\,\Delta\varepsilon_{j+\frac{1}{2}} + (\mathbb{D}f)_{i+\frac{1}{2}}\,\sum_{j=1}^{N} g_{i+\frac{1}{2},j+\frac{1}{2}}\,f_{j+\frac{1}{2}}\,\Delta\varepsilon_{j+\frac{1}{2}}\;,\quad 1 \;\le\; i \;\le\; N \tag{3.20}$$

Like for the previous scheme we can write

$$\widetilde{E}_{i+\frac{1}{2}} \;=\; -\,\sum_{j=1}^{N} g_{i+\frac{1}{2},j+\frac{1}{2}}\,(\mathbb{D}f)_{j+\frac{1}{2}}\,\Delta\varepsilon_{j+\frac{1}{2}} \quad\text{and}\quad \widetilde{D}_{i+\frac{1}{2}} \;=\; \sum_{j=1}^{N} g_{i+\frac{1}{2},j+\frac{1}{2}}\,f_{j+\frac{1}{2}}\,\Delta\varepsilon_{j+\frac{1}{2}}\,,$$

thus

$$\widetilde{K}_{i+\frac{1}{2}} \;=\; \widetilde{E}_{i+\frac{1}{2}}\,f_{i+\frac{1}{2}} \;+\; \widetilde{D}_{i+\frac{1}{2}}\,\frac{f_{i+1} \;-\; f_i}{\Delta\varepsilon_{i+\frac{1}{2}}}\,.$$

Now we have

$$\widetilde{K}_{N+\frac{1}{2}} \;=\; \widetilde{E}_{N+\frac{1}{2}}\,f_{N+\frac{1}{2}} \;+\; \widetilde{D}_{N+\frac{1}{2}}\,\frac{f_{N+1} - f_N}{\Delta\varepsilon_{N+\frac{1}{2}}}\,,$$

as $f_{N+1} = 0$ the flux at the last point reads

$$\widetilde{K}_{N+\frac{1}{2}} \;=\; \widetilde{E}_{N+\frac{1}{2}}\,f_{N+\frac{1}{2}} \;-\; \widetilde{D}_{N+\frac{1}{2}}\,\frac{f_N}{\Delta\varepsilon_{N+\frac{1}{2}}}\,.$$

We enforce $\widetilde{K}_{N+\frac{1}{2}}$ to vanish (to impose $\dfrac{df_{N+1}}{dt} = 0$). Thus we get the relation

$$\widetilde{E}_{N+\frac{1}{2}}\,f_{N+\frac{1}{2}} \;-\; \widetilde{D}_{N+\frac{1}{2}}\,\frac{f_N}{\Delta\varepsilon_{N+\frac{1}{2}}} \;=\; 0\,. \tag{3.21}$$

Thanks to (3.21) the diffusion becomes

$$\widetilde{D}_{i+\frac{1}{2}} = \sum_{j=1}^{N-1} g_{i+\frac{1}{2},j+\frac{1}{2}}\, f_{j+\frac{1}{2}}\, \Delta\varepsilon_{j+\frac{1}{2}} + g_{i+\frac{1}{2},N+\frac{1}{2}}\, \frac{\widetilde{D}_{N+\frac{1}{2}}}{\widetilde{E}_{N+\frac{1}{2}}}\, f_N\,. \tag{3.22}$$

Remark 7. *We integrate by parts the drift term and get*

$$\widetilde{E}_{i+\frac{1}{2}} = \sum_{j=1}^{N} (g_{i+\frac{1}{2},j+\frac{1}{2}} - g_{i+\frac{1}{2},j-\frac{1}{2}})\, f_j\, \Delta\varepsilon_{j+\frac{1}{2}}\,.$$

We assume that $\{g_{i+\frac{1}{2},j+\frac{1}{2}}\}_{1\le i\le N}$ *is an increasing sequence. Then, if the* f_i*'s are positive,* $\widetilde{E}_{i+\frac{1}{2}}$ *is positive too. In other hand, even if the* $f_{i+\frac{1}{2}}$*'s are positive* $\widetilde{D}_{i+\frac{1}{2}}$ *can be negative if* $f_N \neq 0$*.*

The discretization of $f_{i+\frac{1}{2}}$ is still given by the Chang and Cooper average. Therefore we get

$$f_{i+\frac{1}{2}} = (1 - \widetilde{\delta}_{i+\frac{1}{2}}) f_{i+1} + \widetilde{\delta}_{i+\frac{1}{2}}\, f_i\,, \quad \forall\, i;\ 1 \le i \le N-1\,, \tag{3.23}$$

where

$$\widetilde{\delta}_{i+\frac{1}{2}} = h(\widetilde{\alpha}_{i+\frac{1}{2}}) = \frac{1}{\widetilde{\alpha}_{i+\frac{1}{2}}} - \frac{1}{(\exp(\widetilde{\alpha}_{i+\frac{1}{2}}) - 1)}\,, \quad \forall\, i;\ 1 \le i \le N-1\,, \tag{3.24}$$

and with $\widetilde{\alpha}_{i+\frac{1}{2}} = \dfrac{\widetilde{E}_{i+\frac{1}{2}}}{\widetilde{D}_{i+\frac{1}{2}}} \Delta\varepsilon_{i+\frac{1}{2}}$. Thanks to the boundary condition (3.21) we can extend the definition to $N+\frac{1}{2}$ by setting $\widetilde{\delta}_{N+\frac{1}{2}} = 1/\widetilde{\alpha}_{N+\frac{1}{2}}$ (see [15]). Now $\widetilde{\alpha}_{i+\frac{1}{2}}$ reads

$$\widetilde{\alpha}_{i+\frac{1}{2}} = \left(\frac{-\sum_{j=1}^{N} g_{i+\frac{1}{2},j+\frac{1}{2}}\, (\mathbb{D}f)_{j+\frac{1}{2}}\, \Delta\varepsilon_{j+\frac{1}{2}}}{\sum_{j=1}^{N-1} g_{i+\frac{1}{2},j+\frac{1}{2}}\, ((1-\widetilde{\delta}_{j+\frac{1}{2}}) f_{j+1} + \widetilde{\delta}_{j+\frac{1}{2}} f_j)\, \Delta\varepsilon_{j+\frac{1}{2}} + g_{i+\frac{1}{2},N+\frac{1}{2}}\, \dfrac{\widetilde{D}_{N+\frac{1}{2}}}{\widetilde{E}_{N+\frac{1}{2}}}\, f_N} \right) \Delta\varepsilon_{i+\frac{1}{2}}\,,$$

As in the previous scheme we suppose that $\widetilde{\delta}$ is the N-dimensional column vector with components $\{\widetilde{\delta}_{i+\frac{1}{2}}\}_{1\le i\le N}$. Thus we can write $\widetilde{\delta}_{i+\frac{1}{2}}$ as a function of f and introduce $\{\widetilde{\mathcal{H}}_{i+\frac{1}{2}}\}_{1\le i\le N}$ such that

$$\widetilde{\delta}_{i+\frac{1}{2}}(f) = \widetilde{\mathcal{H}}_{i+\frac{1}{2}}(\widetilde{\delta}_{\frac{3}{2}}, ..., \widetilde{\delta}_{k+\frac{1}{2}}, ..., \widetilde{\delta}_{N+\frac{1}{2}}, f_1, ..., f_k, ... f_N)\,,$$

in other words, to compute $\widetilde{\delta}$, the system to be solved is

$$\widetilde{\delta}(f) = \widetilde{\mathcal{H}}(\widetilde{\delta}(f), f)\,, \tag{3.25}$$

where f is a distribution function that has density $\overline{\boldsymbol{\rho}}$ and energy $\overline{\boldsymbol{\rho}\mathbf{E}}$ and $\widetilde{\mathcal{H}}$ is a N-dimensional vector valued function. Here the problem rises from the fact that $\{\widetilde{D}_{i+\frac{1}{2}}\}_{1\le i\le N}$ can be undefined even if f is properly defined. We'll specify this more clearly in the case where potential are Coulombian or Maxwellian.

Proposition 2. *The flux* $\widetilde{K}_{i+\frac{1}{2}}$ *satisfies the relation*

$$\widetilde{K}_{i+\frac{1}{2}} = \widetilde{A}_{i+\frac{1}{2}} f_{i+1} - \widetilde{B}_{i+\frac{1}{2}} f_i \tag{3.26}$$

where

$$\widetilde{B}_{i+\frac{1}{2}} = u(\widetilde{\alpha}_{i+\frac{1}{2}}) \frac{\widetilde{D}_{i+\frac{1}{2}}}{\Delta\varepsilon_{i+\frac{1}{2}}}, \tag{3.27}$$

and

$$\widetilde{A}_{i+\frac{1}{2}} = v(\widetilde{\alpha}_{i+\frac{1}{2}}) \frac{\widetilde{D}_{i+\frac{1}{2}}}{\Delta\varepsilon_{i+\frac{1}{2}}}, \tag{3.28}$$

with $\widetilde{\alpha}_{i+\frac{1}{2}} = \dfrac{\widetilde{E}_{i+\frac{1}{2}}}{\widetilde{D}_{i+\frac{1}{2}}} \Delta\varepsilon_{i+\frac{1}{2}}$, $u(\alpha) = \dfrac{\alpha}{\exp(\alpha)-1}$ *and* $v(\alpha) = \dfrac{\alpha \exp(\alpha)}{\exp(\alpha)-1}$.

The proof is checked as in Proposition 1.

Summary. To summarize we identify as scheme $\mathcal{S}_2$ the system (3.19) with boundary conditions (3.21) where the numerical fluxes are

$\widetilde{K}_{i+\frac{1}{2}} = \widetilde{E}_{i+\frac{1}{2}} f_{i+\frac{1}{2}} + \widetilde{D}_{i+\frac{1}{2}} (f_{i+1} - f_i)/\Delta\varepsilon_{i+\frac{1}{2}}$ with $\widetilde{E}_{i+\frac{1}{2}} = -\sum_{j=1}^{N} g_{i+\frac{1}{2},j+\frac{1}{2}} (\mathbb{D}f)_{j+\frac{1}{2}} \Delta\varepsilon_{j+\frac{1}{2}}$ and

$\widetilde{D}_{i+\frac{1}{2}} = \sum_{j=1}^{N-1} g_{i+\frac{1}{2},j+\frac{1}{2}} f_{j+\frac{1}{2}} \Delta\varepsilon_{j+\frac{1}{2}} + g_{i+\frac{1}{2},N+\frac{1}{2}} f_N \widetilde{D}_{N+\frac{1}{2}} / \widetilde{E}_{N+\frac{1}{2}}$. Here $f_{i+\frac{1}{2}}$ is given by the Chang and Cooper average (3.23)-(3.24).

Remark 8. *Comparing formally the drift and diffusion functional discretized respectively by* $\mathcal{S}_1$ *and* $\mathcal{S}_2$ *we get for* $1 \leq i \leq N-1$

$$\widetilde{E}_{i+\frac{1}{2}} = E_{i+\frac{1}{2}} + f_N \, g_{i+\frac{1}{2},N+\frac{1}{2}} \quad \text{and} \quad \widetilde{D}_{i+\frac{1}{2}} = D_{i+\frac{1}{2}} + \Delta\varepsilon_{N+\frac{1}{2}} \, f_{N+\frac{1}{2}} \, g_{i+\frac{1}{2},N+\frac{1}{2}} \, . \tag{3.29}$$

Remark 9. *Langdon [24] has previously developed a scheme for Coulombian interactions. This scheme is nothing else than* $\mathcal{S}_2$ *reduced to the Coulombic potentials.*

3.1.3 A particular case: the Maxwellian interactions

First we focus on the expression of the collision terms in the case where potentials are Maxwellian. The diffusion-drift ratio is constant and consequently the Chang and Cooper relation is reduced to a non-linear scalar equation. As a result we'll show that for the scheme $\mathcal{S}_2$ the diffusion term can be undefined.

For Maxwellian interactions we have $g_{i+\frac{1}{2},j+\frac{1}{2}} = \varepsilon_{i+\frac{1}{2}}^{\frac{3}{2}} \varepsilon_{j+\frac{1}{2}}^{\frac{3}{2}}$. Then, $E_{i+\frac{1}{2}}$ can be write by factorizing the term f_i

$$E_{i+\frac{1}{2}} = \varepsilon_{i+\frac{1}{2}}^{\frac{3}{2}} \sum_{j=1}^{N-1} (\varepsilon_{j+\frac{1}{2}}^{\frac{3}{2}} - \varepsilon_{j-\frac{1}{2}}^{\frac{3}{2}}) f_j - f_N \varepsilon_{N-\frac{1}{2}}^{\frac{3}{2}} \varepsilon_{i+\frac{1}{2}}^{\frac{3}{2}}$$

and we get for the diffusion term

$$D_{i+\frac{1}{2}} \;=\; \varepsilon_{i+\frac{1}{2}}^{\frac{3}{2}} \sum_{j=1}^{N-1} \varepsilon_{j+\frac{1}{2}}^{\frac{3}{2}}\, f_{j+\frac{1}{2}}\, \Delta\varepsilon_{j+\frac{1}{2}}\,.$$

Now we turn to the scheme $\mathcal{S}_2$ and obtain in the Maxwellian case, assuming that f is a distribution function that has density $\overline{\boldsymbol{\rho}}$ and energy $\overline{\boldsymbol{\rho}\mathbf{E}}$ and that (3.21) holds,

$$\begin{cases} \widetilde{E}_{i+\frac{1}{2}} \;=\; \varepsilon_{i+\frac{1}{2}}^{\frac{3}{2}} \displaystyle\sum_{j=1}^{N} (\varepsilon_{j+\frac{1}{2}}^{\frac{3}{2}} \;-\; \varepsilon_{j-\frac{1}{2}}^{\frac{3}{2}})\, f_j \;=\; \frac{3}{2}\, \varepsilon_{i+\frac{1}{2}}^{\frac{3}{2}} \sum_{j=1}^{N} \sqrt{\varepsilon_j}\, \Delta\varepsilon_j\, f_j \;=\; \frac{3}{2}\, \varepsilon_{i+\frac{1}{2}}^{\frac{3}{2}}\, \overline{\boldsymbol{\rho}}\,, \\ \widetilde{D}_{i+\frac{1}{2}} \;=\; \varepsilon_{i+\frac{1}{2}}^{\frac{3}{2}} \displaystyle\sum_{j=1}^{N} \Delta\varepsilon_{j+\frac{1}{2}}\, \varepsilon_{j+\frac{1}{2}}^{\frac{3}{2}}\, f_{j+\frac{1}{2}} = \varepsilon_{i+\frac{1}{2}}^{\frac{3}{2}} \sum_{j=1}^{N-1} \Delta\varepsilon_{j+\frac{1}{2}}\, \varepsilon_{j+\frac{1}{2}}^{\frac{3}{2}}\, f_{j+\frac{1}{2}} + \varepsilon_{i+\frac{1}{2}}^{\frac{3}{2}} \varepsilon_{N+\frac{1}{2}}^{\frac{3}{2}}\; f_N \frac{\widetilde{D}_{N+\frac{1}{2}}}{\widetilde{E}_{N+\frac{1}{2}}}\,, \end{cases} \tag{3.30}$$

thus

$$\widetilde{E}_{i+\frac{1}{2}} \;=\; E_{i+\frac{1}{2}} \;+\; \varepsilon_{i+\frac{1}{2}}^{\frac{3}{2}}\, \varepsilon_{N+\frac{1}{2}}^{\frac{3}{2}}\, f_N \quad \text{and} \quad \widetilde{D}_{i+\frac{1}{2}} \;=\; D_{i+\frac{1}{2}} \;+\; \varepsilon_{i+\frac{1}{2}}^{\frac{3}{2}}\, \varepsilon_{N+\frac{1}{2}}^{\frac{3}{2}} f_N \frac{\widetilde{D}_{N+\frac{1}{2}}}{\widetilde{E}_{N+\frac{1}{2}}}\,, \tag{3.31}$$

as expected from (3.29).

Remark 10. *We have $D_{i+\frac{1}{2}}/E_{i+\frac{1}{2}} = C$ for i; $1 \le i \le N-1$ and $\widetilde{D}_{i+\frac{1}{2}}/\widetilde{E}_{i+\frac{1}{2}} = \widetilde{C}$ for i; $1 \le i \le N$. Following (3.21) and (3.31) we can write*

$$\frac{\widetilde{D}_{i+\frac{1}{2}}}{\widetilde{E}_{i+\frac{1}{2}}} \;=\; \left(\frac{D_{i+\frac{1}{2}} \dfrac{\widetilde{E}_{N+\frac{1}{2}}}{\widetilde{D}_{N+\frac{1}{2}}} \;+\; \varepsilon_{i+\frac{1}{2}}^{\frac{3}{2}}\, \varepsilon_{N+\frac{1}{2}}^{\frac{3}{2}}\, f_N}{E_{i+\frac{1}{2}} \;+\; \varepsilon_{i+\frac{1}{2}}^{\frac{3}{2}}\, \varepsilon_{N+\frac{1}{2}}^{\frac{3}{2}}\, f_N} \right) \frac{\widetilde{D}_{N+\frac{1}{2}}}{\widetilde{E}_{N+\frac{1}{2}}}\,.$$

As $(\widetilde{D}_{i+\frac{1}{2}}/\widetilde{E}_{i+\frac{1}{2}}) = (\widetilde{D}_{N+\frac{1}{2}}/\widetilde{E}_{N+\frac{1}{2}})$ we get $C = \widetilde{C}$.

Now, we are interested by the expression of $\delta_{i+\frac{1}{2}}$ in the case where potentials are Maxwellian. Recall that with scheme $\mathcal{S}_1$ we get $\alpha_{i+\frac{1}{2}} = E_{i+\frac{1}{2}}/D_{i+\frac{1}{2}}\, \Delta\varepsilon_{i+\frac{1}{2}}$ and $h(\alpha) = 1/\alpha - 1/(\exp(\alpha) - 1)$. Thus, thanks to (3.30) and (3.31) we can write

$$\alpha_{i+\frac{1}{2}} \;=\; \frac{\frac{3}{2}\overline{\boldsymbol{\rho}} - f_N \varepsilon_{N+\frac{1}{2}}^{\frac{3}{2}}}{\displaystyle\sum_{j=1}^{N-1} \varepsilon_{j+\frac{1}{2}}^{\frac{3}{2}}\, \Delta\varepsilon_{j+\frac{1}{2}}\, f_{j+\frac{1}{2}}}\, \Delta\varepsilon_{i+\frac{1}{2}}\,.$$

We set $d(f) = \displaystyle\sum_{j=1}^{N-1} \varepsilon_{j+\frac{1}{2}}^{\frac{3}{2}}\, \Delta\varepsilon_{j+\frac{1}{2}}\, f_{j+\frac{1}{2}}$ therefore as $f_{i+\frac{1}{2}}$ is given by the Chang and Cooper average

$$d(f) \;=\; \sum_{j=1}^{N-1} \Delta\varepsilon_{j+\frac{1}{2}}\, \varepsilon_{j+\frac{1}{2}}^{\frac{3}{2}}\, ((1-\delta_{j+\frac{1}{2}})f_{j+1} + \delta_{j+\frac{1}{2}} f_j)\,,$$

and introducing $h(\alpha_{i+\frac{1}{2}})$ we get

$$d(f) \;=\; \sum_{j=1}^{N-1} \Delta\varepsilon_{j+\frac{1}{2}}\, \varepsilon_{j+\frac{1}{2}}^{\frac{3}{2}}\, (h(\frac{\frac{3}{2}\overline{\boldsymbol{\rho}} - f_N \varepsilon_{N+\frac{1}{2}}^{\frac{3}{2}}}{d(f)}\Delta\varepsilon_{j+\frac{1}{2}})\,(f_j - f_{j+1}) \;+\; f_{j+1})\,. \tag{3.32}$$

Note that $d(f)$ doesn't depend on $(i+\frac{1}{2})$. Thus, the Chang and Cooper relation is reduced to a simple scalar non-linear equation. To find $\delta_{i+\frac{1}{2}}$ all we need to do is to solve equation (3.32) and that can be achieved by a Newton method in only $\mathcal{O}(N)$ operations at each step. Moreover if the grid is uniform all the $\alpha_{i+\frac{1}{2}}$ and $\delta_{i+\frac{1}{2}}$ are equal and one can check directly on (3.6) that the scheme reduces to the scheme of Berezin and Pekker [1]

$$\sum_{i=1}^{N} c_i \frac{\partial f_i}{\partial t}\phi_i \;=\; -\sum_{i=1}^{N-1} (\mathbb{D}\phi)_{i+\frac{1}{2}}\, \Delta\varepsilon_{i+\frac{1}{2}} \sum_{j=1}^{N-1} g_{i+\frac{1}{2},j+\frac{1}{2}}\, (f_{i+1}f_j - f_i f_{j+1}),$$

which is entropic as shown in [5].

Following the same way we get for the scheme $\mathcal{S}_2$

$$\widetilde{\alpha}_{i+\frac{1}{2}} \;=\; \frac{\frac{3}{2}\overline{\boldsymbol{\rho}}}{\displaystyle\sum_{j=1}^{N-1} \varepsilon_{j+\frac{1}{2}}^{\frac{3}{2}}\, \Delta\varepsilon_{j+\frac{1}{2}}\, f_{j+\frac{1}{2}} \;+\; \varepsilon_{N+\frac{1}{2}}^{\frac{3}{2}} f_{N+\frac{1}{2}} \Delta\varepsilon_{N+\frac{1}{2}}}\, \Delta\varepsilon_{i+\frac{1}{2}}\,.$$

We denote $\widetilde{d}(f) = \displaystyle\sum_{j=1}^{N-1} \varepsilon_{j+\frac{1}{2}}^{\frac{3}{2}}\, \Delta\varepsilon_{j+\frac{1}{2}}\, f_{j+\frac{1}{2}} + \varepsilon_{N+\frac{1}{2}}^{\frac{3}{2}} f_{N+\frac{1}{2}} \Delta\varepsilon_{N+\frac{1}{2}}$ where $f_{i+\frac{1}{2}}$ is computed by the Chang and Cooper average. Thus

$$\widetilde{d}(f) \;=\; \sum_{j=1}^{N-1} \Delta\varepsilon_{j+\frac{1}{2}}\, \varepsilon_{j+\frac{1}{2}}^{\frac{3}{2}}\, ((1-\widetilde{\delta}_{j+\frac{1}{2}})f_{j+1} + \widetilde{\delta}_{j+\frac{1}{2}} f_j) \;+\; \varepsilon_{N+\frac{1}{2}}^{\frac{3}{2}} f_{N+\frac{1}{2}} \Delta\varepsilon_{N+\frac{1}{2}}\,,$$

so that, according that (3.21) holds and introducing $h(\widetilde{\alpha}_{i+\frac{1}{2}})$, we have

$$\widetilde{d}(f) \;=\; \sum_{j=1}^{N-1} \Delta\varepsilon_{j+\frac{1}{2}}\, \varepsilon_{j+\frac{1}{2}}^{\frac{3}{2}}\, (h(\frac{\frac{3}{2}\overline{\boldsymbol{\rho}}}{\widetilde{d}(f)}\Delta\varepsilon_{j+\frac{1}{2}})\,(f_j - f_{j+1}) \;+\; f_{j+1}) \;+\; \varepsilon_{N+\frac{1}{2}}^{\frac{3}{2}} f_N \frac{\widetilde{d}(f)}{\frac{3}{2}\overline{\boldsymbol{\rho}}}\,. \tag{3.33}$$

Hence if $\dfrac{3}{2}\overline{\boldsymbol{\rho}} = \varepsilon_{N+\frac{1}{2}}^{\frac{3}{2}} f_N$ the term $\widetilde{d}(f)$ is undefined. As the $f_{i+\frac{1}{2}}$'s are positive we remark that $\widetilde{d}(f)$ and

$$Y = (\frac{3}{2}\overline{\boldsymbol{\rho}} - \varepsilon_{N+\frac{1}{2}}^{\frac{3}{2}} f_N) \tag{3.34}$$

have the same sign. To find $\widetilde{\delta}_{i+\frac{1}{2}}$ we have to solve the non-linear equation (3.33).

3.1.4 A particular case: the Coulombian interactions

In a first part we give the expression of drift and diffusion in the case where the potentials are Coulombian. For $\mathcal{S}_2$ we find Langdon's scheme as planned. We point out that $\{\widetilde{D}_{i+\frac{1}{2}}\}_{1\leq i\leq N}$ (diffusion term computed by the scheme $\mathcal{S}_2$) can be undefined. After that we carry on in offering an equivalent form of (3.13) (respectively (3.25)) to compute the interpolant of the Chang and Cooper average as in [14]. An advantage is that this form requires less operation to implement.

For Coulombian interactions we get $g_{i+\frac{1}{2},j+\frac{1}{2}} = g(\varepsilon_{i+\frac{1}{2}},\varepsilon_{j+\frac{1}{2}}) = \min(\varepsilon^{\frac{3}{2}}_{i+\frac{1}{2}},\varepsilon^{\frac{3}{2}}_{j+\frac{1}{2}})$. This relation is inserted in (3.9) to obtain the drift term discretized by the scheme $\mathcal{S}_1$

$$E_{i+\frac{1}{2}} \;=\; -\sum_{j=1}^{N-1} \Delta\varepsilon_{j+\frac{1}{2}}\; \min(\varepsilon^{\frac{3}{2}}_{i+\frac{1}{2}},\varepsilon^{\frac{3}{2}}_{j+\frac{1}{2}})\,\frac{f_{j+1}\,-\,f_j}{\Delta\varepsilon_{j+\frac{1}{2}}} \qquad \text{for } i\;;\; 1\leq i\leq N-1\,,$$

by factorizing the term f_i we have

$$E_{i+\frac{1}{2}} \;=\; \sum_{j=1}^{N-1} (\min(\varepsilon^{\frac{3}{2}}_{i+\frac{1}{2}},\varepsilon^{\frac{3}{2}}_{j+\frac{1}{2}}) \;-\; \min(\varepsilon^{\frac{3}{2}}_{i+\frac{1}{2}},\varepsilon^{\frac{3}{2}}_{j-\frac{1}{2}}))\, f_j \;-\; f_N\, \min(\varepsilon^{\frac{3}{2}}_{i+\frac{1}{2}},\varepsilon^{\frac{3}{2}}_{N-\frac{1}{2}})\,. \tag{3.35}$$

Now $\{\varepsilon_{i+\frac{1}{2}}\}_{1\leq i\leq N}$ is an increasing sequence, therefore

$$E_{i+\frac{1}{2}} \;=\; \sum_{j=1}^{i} (\varepsilon^{\frac{3}{2}}_{j+\frac{1}{2}} \;-\; \varepsilon^{\frac{3}{2}}_{j-\frac{1}{2}})\, f_j \;-\; f_N\, \min(\varepsilon^{\frac{3}{2}}_{i+\frac{1}{2}},\varepsilon^{\frac{3}{2}}_{N-\frac{1}{2}}) \;\;\forall\, i;\;\; 1\;\leq\; i\;\leq\; N-1\,. \tag{3.36}$$

In the same way, we get for the diffusion term

$$D_{i+\frac{1}{2}} \;=\; \sum_{j=1}^{N-1} \Delta\varepsilon_{j+\frac{1}{2}}\; \min(\varepsilon^{\frac{3}{2}}_{i+\frac{1}{2}},\varepsilon^{\frac{3}{2}}_{j+\frac{1}{2}})\, f_{j+\frac{1}{2}} \;\;\forall\, i;\;\; 1\;\leq\; i\;\leq\; N-1\;, \tag{3.37}$$

or

$$\begin{cases} D_{i+\frac{1}{2}} \;=\; \displaystyle\sum_{j=1}^{i} \Delta\varepsilon_{j+\frac{1}{2}}\, f_{j+\frac{1}{2}}\, \varepsilon^{\frac{3}{2}}_{j+\frac{1}{2}} \;+\; \varepsilon^{\frac{3}{2}}_{i+\frac{1}{2}} \sum_{j=i+1}^{N-1} \Delta\varepsilon_{j+\frac{1}{2}}\, f_{j+\frac{1}{2}}\,, \quad 1\leq i\leq N-2 \\ D_{N-\frac{1}{2}} \;=\; \displaystyle\sum_{j=1}^{N-1} \Delta\varepsilon_{j+\frac{1}{2}} f_{j+\frac{1}{2}} \varepsilon^{\frac{3}{2}}_{j+\frac{1}{2}}\,. \end{cases} \tag{3.38}$$

By construction (3.36) and (3.38) are consistent approximations of drift $E(f)$ and diffusion $D(f)$ when their integration domain is reduced to a bounded domain $[0,\mathcal{E}]$ in the variable ε (see equations (2.13) and (2.14)).
Now, the collision terms computed by the scheme $\mathcal{S}_2$ read in the case of Coulombian interactions

$$\widetilde{E}_{i+\frac{1}{2}} = -\sum_{j=1}^{N} \min(\varepsilon^{\frac{3}{2}}_{i+\frac{1}{2}},\varepsilon^{\frac{3}{2}}_{j+\frac{1}{2}})(f_{j+1}-f_j) = \sum_{j=1}^{N} (\min(\varepsilon^{\frac{3}{2}}_{i+\frac{1}{2}},\varepsilon^{\frac{3}{2}}_{j+\frac{1}{2}}) - \min(\varepsilon^{\frac{3}{2}}_{i+\frac{1}{2}},\varepsilon^{\frac{3}{2}}_{j-\frac{1}{2}}))f_j\,, \tag{3.39}$$

$$\widetilde{E}_{i+\frac{1}{2}} = \sum_{j=1}^{i} (\varepsilon_{j+\frac{1}{2}}^{\frac{3}{2}} - \varepsilon_{j-\frac{1}{2}}^{\frac{3}{2}})\, f_j \quad \text{and} \quad \widetilde{E}_{N-\frac{1}{2}} = \frac{3}{2}\overline{\boldsymbol{\rho}} - (\varepsilon_{N+\frac{1}{2}}^{\frac{3}{2}} - \varepsilon_{N-\frac{1}{2}}^{\frac{3}{2}})\, f_N, \tag{3.40}$$

and

$$\widetilde{D}_{i+\frac{1}{2}} = \sum_{j=1}^{N-1} \Delta\varepsilon_{j+\frac{1}{2}} \min(\varepsilon_{i+\frac{1}{2}}^{\frac{3}{2}}, \varepsilon_{j+\frac{1}{2}}^{\frac{3}{2}}) f_{j+\frac{1}{2}} \; + \; \Delta\varepsilon_{N+\frac{1}{2}}\, f_{N+\frac{1}{2}}\, \varepsilon_{i+\frac{1}{2}}^{\frac{3}{2}}\,, \tag{3.41}$$

or,

$$\begin{cases} \widetilde{D}_{i+\frac{1}{2}} = \displaystyle\sum_{j=1}^{i} \varepsilon_{j+\frac{1}{2}}^{\frac{3}{2}} f_{j+\frac{1}{2}} \Delta\varepsilon_{j+\frac{1}{2}} + \varepsilon_{i+\frac{1}{2}}^{\frac{3}{2}} \sum_{j=i+1}^{N-1} f_{j+\frac{1}{2}} \Delta\varepsilon_{j+\frac{1}{2}} + \Delta\varepsilon_{N+\frac{1}{2}}\, f_{N+\frac{1}{2}}\, \varepsilon_{i+\frac{1}{2}}^{\frac{3}{2}}, \qquad 1 \le i \le N-2 \\ \widetilde{D}_{i+\frac{1}{2}} = \displaystyle\sum_{j=1}^{N-1} \varepsilon_{j+\frac{1}{2}}^{\frac{3}{2}} f_{j+\frac{1}{2}} \Delta\varepsilon_{j+\frac{1}{2}} + \Delta\varepsilon_{N+\frac{1}{2}} f_{N+\frac{1}{2}} \varepsilon_{i+\frac{1}{2}}^{\frac{3}{2}}, \qquad i = N-1 \text{ or } i = N\,. \end{cases}$$

Note that contrary to scheme $\mathcal{S}_1$ the drift and diffusion terms computed by scheme $\mathcal{S}_2$ are not consistent approximations of $E(f)$ and $D(f)$ when their integration domain is reduced to a bounded domain $[0, \mathcal{E}]$ in the variable ε (see equations (2.13) and (2.14)). To compare to the Langdon scheme, we prefer to write $E_{i+\frac{1}{2}}$, $D_{i+\frac{1}{2}}$, $\widetilde{E}_{i+\frac{1}{2}}$ and $\widetilde{D}_{i+\frac{1}{2}}$ with the velocity variable when it is necessary. Thanks to (3.2) we have

$$E_{i+\frac{1}{2}} \; = \; 3 \sum_{j=1}^{i} (v^2 \Delta v)_j\, f_j \; - \; f_N \, \min(\varepsilon_{i+\frac{1}{2}}^{\frac{3}{2}}, \varepsilon_{N-\frac{1}{2}}^{\frac{3}{2}})\,. \tag{3.42}$$

Seeing $\varepsilon_{\frac{1}{2}} = 0$ we can write $\varepsilon_{i+\frac{1}{2}}^{\frac{3}{2}} = \displaystyle\sum_{j=1}^{i} (\varepsilon_{j+\frac{1}{2}}^{\frac{3}{2}} - \varepsilon_{j-\frac{1}{2}}^{\frac{3}{2}})$ and from (3.2) we get $\varepsilon_{i+\frac{1}{2}}^{\frac{3}{2}} = 3 \displaystyle\sum_{j=1}^{i} (v^2 \Delta v)_j$.
So

$$\begin{cases} D_{i+\frac{1}{2}} \; = \; 3 \displaystyle\sum_{j=1}^{i} \Delta\varepsilon_{j+\frac{1}{2}}\, f_{j+\frac{1}{2}} \sum_{k=1}^{j} (v^2 \Delta v)_k \; + \; 3 \sum_{j=i+1}^{N-1} \Delta\varepsilon_{j+\frac{1}{2}}\, f_{j+\frac{1}{2}} \sum_{k=1}^{i} (v^2 \Delta v)_k, \qquad 1 \le i \le N-2 \\ D_{N-\frac{1}{2}} \; = \; 3 \displaystyle\sum_{j=1}^{N-1} \Delta\varepsilon_{j+\frac{1}{2}}\, f_{j+\frac{1}{2}} \sum_{k=1}^{j} (v^2 \Delta v)_k \end{cases}$$

By factorizing the term $(v^2 \Delta v)_k$ we write

$$D_{i+\frac{1}{2}} \; = \; 3 \sum_{j=1}^{i} (v^2 \Delta v)_j \sum_{k=j}^{N-1} f_{k+\frac{1}{2}}\, \Delta\varepsilon_{k+\frac{1}{2}} \quad 1 \le i \le N-1\,. \tag{3.43}$$

For the scheme $\mathcal{S}_2$ we obviously find $\widetilde{E}_{i+\frac{1}{2}} = 3 \displaystyle\sum_{j=1}^{i} (v^2 \Delta v)_j\, f_j$ and $\widetilde{D}_{i+\frac{1}{2}} = 3 \displaystyle\sum_{j=1}^{i} (v^2 \Delta v)_j \sum_{k=j}^{N} f_{k+\frac{1}{2}} \Delta\varepsilon_{k+\frac{1}{2}}$ as in Langdon's scheme [24]. That is, following (3.42) and (3.43), for $1 \le i \le N-1$

$$E_{i+\frac{1}{2}} = \widetilde{E}_{i+\frac{1}{2}} - f_N \min(\varepsilon^{\frac{3}{2}}_{i+\frac{1}{2}}, \varepsilon^{\frac{3}{2}}_{N-\frac{1}{2}}) \text{ and } D_{i+\frac{1}{2}} = \widetilde{D}_{i+\frac{1}{2}} - f_{N+\frac{1}{2}}\, \varepsilon^{\frac{3}{2}}_{i+\frac{1}{2}}\, \Delta\varepsilon_{N+\frac{1}{2}}\,. \quad (3.44)$$

Remark 11. *In the Coulombian case, thanks to (3.35), (3.39), (3.37), (3.41) and (3.21) we can write*

$$\widetilde{E}_{N+\frac{1}{2}} = E_{N-\frac{1}{2}} + \varepsilon^{\frac{3}{2}}_{N+\frac{1}{2}}\, f_N \text{ and } \widetilde{D}_{N+\frac{1}{2}} = D_{N-\frac{1}{2}} + \varepsilon^{\frac{3}{2}}_{N+\frac{1}{2}} \Delta\varepsilon_{N+\frac{1}{2}} f_{N+\frac{1}{2}} = D_{N-\frac{1}{2}} + \varepsilon^{\frac{3}{2}}_{N+\frac{1}{2}}\, f_N\, \frac{\widetilde{D}_{N+\frac{1}{2}}}{\widetilde{E}_{N+\frac{1}{2}}}\,,$$

consequently

$$\frac{\widetilde{D}_{N+\frac{1}{2}}}{\widetilde{E}_{N+\frac{1}{2}}} = \frac{\widetilde{D}_{N+\frac{1}{2}}}{\widetilde{E}_{N+\frac{1}{2}}}\, \frac{D_{N-\frac{1}{2}}\, \dfrac{\widetilde{E}_{N+\frac{1}{2}}}{\widetilde{D}_{N+\frac{1}{2}}} + \varepsilon^{\frac{3}{2}}_{N+\frac{1}{2}}\, f_N}{E_{N-\frac{1}{2}} + \varepsilon^{\frac{3}{2}}_{N+\frac{1}{2}}\, f_N}\,.$$

This leads to the relation

$$\frac{\widetilde{D}_{N+\frac{1}{2}}}{\widetilde{E}_{N+\frac{1}{2}}} = \frac{D_{N-\frac{1}{2}}}{E_{N-\frac{1}{2}}}\,.$$

Using (3.21) and (3.44) again, we write

$$\widetilde{E}_{N-\frac{1}{2}} = E_{N-\frac{1}{2}} + \varepsilon^{\frac{3}{2}}_{N-\frac{1}{2}}\, f_N \text{ and } \widetilde{D}_{N-\frac{1}{2}} = D_{N-\frac{1}{2}} + \varepsilon^{\frac{3}{2}}_{N-\frac{1}{2}}\, f_N\, \frac{\widetilde{D}_{N+\frac{1}{2}}}{\widetilde{E}_{N+\frac{1}{2}}}\,,$$

hence

$$\frac{\widetilde{D}_{N+\frac{1}{2}}}{\widetilde{E}_{N+\frac{1}{2}}} = \frac{\widetilde{D}_{N-\frac{1}{2}}}{\widetilde{E}_{N-\frac{1}{2}}}\,. \quad (3.45)$$

Decoster [14] shows that the relation (3.45) is sufficient to conserve energy and imposes it as boundary condition. Finally this is not necessary in view that (3.45) is held in the discretization.

Here the diffusion term discretized by $\mathcal{S}_2$ reads, thanks to (3.21) and (3.45)

$$\widetilde{D}_{i+\frac{1}{2}} = \sum_{j=1}^{N-1} \Delta\varepsilon_{j+\frac{1}{2}} \min(\varepsilon^{\frac{3}{2}}_{i+\frac{1}{2}}, \varepsilon^{\frac{3}{2}}_{j+\frac{1}{2}}) f_{j+\frac{1}{2}} + \varepsilon^{\frac{3}{2}}_{i+\frac{1}{2}}\, f_N\, \frac{\widetilde{D}_{N-\frac{1}{2}}}{\widetilde{E}_{N-\frac{1}{2}}}\,, \quad (3.46)$$

therefore

$$\widetilde{D}_{N-\frac{1}{2}} = \sum_{j=1}^{N-1} \Delta\varepsilon_{j+\frac{1}{2}}\, \varepsilon^{\frac{3}{2}}_{j+\frac{1}{2}}\, f_{j+\frac{1}{2}} + \varepsilon^{\frac{3}{2}}_{N-\frac{1}{2}}\, f_N\, \frac{\widetilde{D}_{N-\frac{1}{2}}}{\widetilde{E}_{N-\frac{1}{2}}}\,. \quad (3.47)$$

Now if $f_N \varepsilon_{N-\frac{1}{2}}^{\frac{3}{2}} = \widetilde{E}_{N-\frac{1}{2}}$ the term $\widetilde{D}_{N-\frac{1}{2}}$ is undefined. As the $f_{i+\frac{1}{2}}$'s are positive the expression $\sum_{j=1}^{N-1} \Delta\varepsilon_{j+\frac{1}{2}} \varepsilon_{j+\frac{1}{2}}^{\frac{3}{2}} f_{j+\frac{1}{2}}$ remains positive and thus $\widetilde{D}_{N-\frac{1}{2}}$ and

$$Y = (\widetilde{E}_{N-\frac{1}{2}} - f_N \varepsilon_{N-\frac{1}{2}}^{\frac{3}{2}}) = \frac{3}{2}\overline{\boldsymbol{\rho}} - f_N \varepsilon_{N+\frac{1}{2}}^{\frac{3}{2}} \tag{3.48}$$

have the same sign. Because the $\widetilde{E}_{i+\frac{1}{2}}$'s are positive from (3.45) we get that $\widetilde{D}_{N-\frac{1}{2}}$ and $\widetilde{D}_{N+\frac{1}{2}}$ have the same sign. Consequently, thanks to (3.41) and (3.21), if $\widetilde{D}_{N-\frac{1}{2}}$ is positive all the $\widetilde{D}_{i+\frac{1}{2}}$'s are positive. The positivity condition $f_N \varepsilon_{N-\frac{1}{2}}^{\frac{3}{2}} < \widetilde{E}_{N-\frac{1}{2}}$ is the same condition given by Decoster to avoid what he calls "abnormal discretization" [14]. Note that if $\widetilde{D}_{N-\frac{1}{2}}$ is undefined the system (3.25) doesn't have any solution.

In this particular "Coulombian" case an other way to compute the Chang and Cooper average interpolant $\delta_{i+\frac{1}{2}}$ is to consider equation (2.18) as in [14] in place of system (3.13). Thus, we have to solve an elliptic equation, namely

$$-\frac{2}{3}\frac{\partial}{\partial\varepsilon}\frac{1}{\sqrt{\varepsilon}}\frac{\partial}{\partial\varepsilon}(D(f,\,\varepsilon)) \;=\; f(\varepsilon) \qquad \text{in}\,\mathbb{R}^+\,,$$

with boundary condition $D(f,0) = 0$. The solution is $D(f) = \int_0^{\varepsilon} \varepsilon'^{\frac{3}{2}} f(\varepsilon')d\varepsilon' + \int_{\varepsilon}^{\infty} f(\varepsilon')d\varepsilon'$. The non-linear system (3.13) which solution is $\delta(f)$ is then replaced by a non-linear system on $D(f)$.

Proposition 3. *The diffusion term discretized by $\mathcal{S}_1$ is solution of the following system*

$$\mathcal{M}\,D \;=\; f^{\delta} \tag{3.49}$$

where D (respectively f^{δ}) is the $(N-1)$-dimensional column vector with components $\{D_{i+\frac{1}{2}}\}_{1\le i\le N-1}$ (respectively $\{\Delta\varepsilon_{i+\frac{1}{2}} f_{i+\frac{1}{2}}\}_{1\le i\le N-1}$) and $\mathcal{M}$ is the $(N-1)\times(N-1)$ tridiagonal matrix with components

$$\mathcal{M}_{1,2} \;=\; -(\varepsilon_{\frac{5}{2}}^{\frac{3}{2}} - \varepsilon_{\frac{3}{2}}^{\frac{3}{2}})^{-1}\,,\;\; \mathcal{M}_{1,1} \;=\; -\mathcal{M}_{1,2} + (\varepsilon_{\frac{3}{2}}^{\frac{3}{2}})^{-1}$$

$$\mathcal{M}_{i,i-1} \;=\; -(\varepsilon_{i+\frac{3}{2}}^{\frac{3}{2}} - \varepsilon_{i+\frac{1}{2}}^{\frac{3}{2}})^{-1}\,,\;\; \mathcal{M}_{i,i+1} \;=\; -(\varepsilon_{i+\frac{1}{2}}^{\frac{3}{2}} - \varepsilon_{i-\frac{1}{2}}^{\frac{3}{2}})^{-1}\,,\;\; \mathcal{M}_{i,i} \;=\; -\,(\mathcal{M}_{i,i-1} + \mathcal{M}_{i,i+1})\,,$$

$$\mathcal{M}_{N-\frac{1}{2},N-\frac{3}{2}} \;=\; -\,(\varepsilon_{N-\frac{1}{2}}^{\frac{3}{2}} - \varepsilon_{N-\frac{3}{2}}^{\frac{3}{2}})^{-1}\,,\;\; \mathcal{M}_{N-\frac{1}{2},N-\frac{1}{2}} \;=\; -\,\mathcal{M}_{N-\frac{1}{2},N-\frac{3}{2}}\,,$$

Proof. From (3.38) we get for i; $1 \le i \le N-2$

$$D_{i+\frac{1}{2}} \;=\; \sum_{j=1}^{i} \varepsilon_{j+\frac{1}{2}}^{\frac{3}{2}}\, f_{j+\frac{1}{2}}\, \Delta\varepsilon_{j+\frac{1}{2}} \;+\; \varepsilon_{i+\frac{1}{2}}^{\frac{3}{2}} \sum_{j=i+1}^{N-1} f_{j+\frac{1}{2}}\, \Delta\varepsilon_{j+\frac{1}{2}}\,,$$

and for i; $2 \le i \le N-1$

$$D_{i-\frac{1}{2}} \,=\, \sum_{j=1}^{i-1} \varepsilon_{j+\frac{1}{2}}^{\frac{3}{2}}\, f_{j+\frac{1}{2}}\, \Delta\varepsilon_{j+\frac{1}{2}} \,+\, \varepsilon_{i-\frac{1}{2}}^{\frac{3}{2}} \sum_{j=i}^{N-1} f_{j+\frac{1}{2}}\, \Delta\varepsilon_{j+\frac{1}{2}}\,.$$

By subtracting the second equation from the first one we get

$$D_{i+\frac{1}{2}} \,-\, D_{i-\frac{1}{2}} \,=\, \varepsilon_{i+\frac{1}{2}}^{\frac{3}{2}}\, f_{i+\frac{1}{2}}\, \Delta\varepsilon_{i+\frac{1}{2}} \,+\, (\varepsilon_{i+\frac{1}{2}}^{\frac{3}{2}} - \varepsilon_{i-\frac{1}{2}}^{\frac{3}{2}}) \sum_{j=i+1}^{N-1} f_{j+\frac{1}{2}} \Delta\varepsilon_{j+\frac{1}{2}} \,-\, \varepsilon_{i-\frac{1}{2}}^{\frac{3}{2}} f_{i+\frac{1}{2}} \Delta\varepsilon_{i+\frac{1}{2}}\,.$$

Thus

$$\frac{D_{i+\frac{1}{2}} \,-\, D_{i-\frac{1}{2}}}{\varepsilon_{i+\frac{1}{2}}^{\frac{3}{2}} - \varepsilon_{i-\frac{1}{2}}^{\frac{3}{2}}} \,=\, f_{i+\frac{1}{2}} \Delta\varepsilon_{i+\frac{1}{2}} \,+\, \sum_{j=i+1}^{N-1} f_{j+\frac{1}{2}}\, \Delta\varepsilon_{j+\frac{1}{2}}\,.$$

In the same way we have the relation

$$\frac{D_{i+\frac{3}{2}} \,-\, D_{i+\frac{1}{2}}}{\varepsilon_{i+\frac{3}{2}}^{\frac{3}{2}} - \varepsilon_{i+\frac{1}{2}}^{\frac{3}{2}}} \,=\, f_{i+\frac{3}{2}} \Delta\varepsilon_{i+\frac{3}{2}} \,+\, \sum_{j=i+2}^{N-1} f_{j+\frac{1}{2}} \Delta\varepsilon_{j+\frac{1}{2}}\,,$$

hence

$$\frac{D_{i+\frac{3}{2}} \,-\, D_{i+\frac{1}{2}}}{\varepsilon_{i+\frac{3}{2}}^{\frac{3}{2}} - \varepsilon_{i+\frac{1}{2}}^{\frac{3}{2}}} \,-\, \frac{D_{i+\frac{1}{2}} \,-\, D_{i-\frac{1}{2}}}{\varepsilon_{i+\frac{1}{2}}^{\frac{3}{2}} - \varepsilon_{i-\frac{1}{2}}^{\frac{3}{2}}} \,=\, -\, f_{i+\frac{1}{2}} \Delta\varepsilon_{i+\frac{1}{2}}\,, \quad \forall i\,;\; 2 \le i \le N-2\,. \tag{3.50}$$

Now we have to produce the relationship associated to $i = N-1$ and $i = 2$. First, from (3.37) we get relations

$$D_{N-\frac{1}{2}} \,=\, \sum_{j=1}^{N-1} \varepsilon_{j+\frac{1}{2}}^{\frac{3}{2}}\, f_{j+\frac{1}{2}}\, \Delta\varepsilon_{j+\frac{1}{2}} \text{ and } D_{N-\frac{3}{2}} \,=\, \sum_{j=1}^{N-2} \varepsilon_{j+\frac{1}{2}}^{\frac{3}{2}}\, f_{j+\frac{1}{2}} \Delta\varepsilon_{j+\frac{1}{2}} + \varepsilon_{N-\frac{3}{2}}^{\frac{3}{2}} f_{N-\frac{1}{2}} \Delta\varepsilon_{N-\frac{1}{2}}\,,$$

and so on

$$\frac{D_{N-\frac{1}{2}} \,-\, D_{N-\frac{3}{2}}}{\varepsilon_{N-\frac{1}{2}}^{\frac{3}{2}} - \varepsilon_{N-\frac{3}{2}}^{\frac{3}{2}}} \,=\, f_{N-\frac{1}{2}}\, \Delta\varepsilon_{N-\frac{1}{2}}\,. \tag{3.51}$$

Furthermore

$$D_{\frac{3}{2}} \,=\, \varepsilon_{\frac{3}{2}}^{\frac{3}{2}} f_{\frac{3}{2}}\, \Delta\varepsilon_{\frac{3}{2}} \,+\, \varepsilon_{\frac{3}{2}}^{\frac{3}{2}} \sum_{j=2}^{N-1} f_{j+\frac{1}{2}}\, \Delta\varepsilon_{j+\frac{1}{2}}\,.$$

We sum the $(N-3)$ relations (3.50) from $i = 2$ to $i = N-2$ and (3.51). Therefore

$$D_{\frac{3}{2}} \,=\, \varepsilon_{\frac{3}{2}}^{\frac{3}{2}} f_{\frac{3}{2}}\, \Delta\varepsilon_{\frac{3}{2}} \,+\, \varepsilon_{\frac{3}{2}}^{\frac{3}{2}} \frac{D_{\frac{5}{2}} \,-\, D_{\frac{3}{2}}}{\varepsilon_{\frac{5}{2}}^{\frac{3}{2}} \,-\, \varepsilon_{\frac{3}{2}}^{\frac{3}{2}}}\,.$$

This completes the proof. □

Remark 12. *The system (3.49) can be solved by a Newton method in only $\mathcal{O}(N)$ operations at each step.*

We can make also the following remark

Remark 13. *$\mathcal{M}$ is a M-matrix. Actually, according that $\{\varepsilon_{i+\frac{1}{2}}\}_{1\leq i\leq N-1}$ is an increasing sequence $\mathcal{M}_{i,i-1}$ and $\mathcal{M}_{i,i+1}$ are negative. Furthermore $\mathcal{M}_{i,i} = -(\mathcal{M}_{i,i-1} + \mathcal{M}_{i,i+1})$ then $\mathcal{M}_{i,i} > 0$. As $\mathcal{M}$ is non-singular $\mathcal{M}$ is a M-matrix (that means $(\mathcal{M}^{-1}v, v) > 0; \forall v > 0$). So, if $\{f_{i+\frac{1}{2}}\}_{1\leq i\leq N-1}$ is positive the $D_{i+\frac{1}{2}}$'s remain positive. This confirms Remark 6. We'll see latter that this property is unavoidable to guarantee the positivity of the scheme $\mathcal{S}_1$.*

We turn now to the scheme $\mathcal{S}_2$ (always in the particular case where potential are Coulombian) and produce the following proposition.

Proposition 4. *Let f a distribution function that has density $\overline{\boldsymbol{\rho}}$ and energy $\overline{\boldsymbol{\rho}\mathbf{E}}$. We assume that $f_N\varepsilon_{N-\frac{1}{2}}^{\frac{3}{2}}/\widetilde{E}_{N-\frac{1}{2}} \neq 1$. The diffusion term discretized by $\mathcal{S}_2$ is solution of the following system*

$$\widetilde{\mathcal{M}}\,\widetilde{D} \,=\, \widetilde{f^{\delta}}$$

where $\widetilde{D}$ (respectively $\widetilde{f^{\delta}}$) is the $(N-1)$-dimensional column vector with components $\{\widetilde{D}_{i+\frac{1}{2}}\}_{1\leq i\leq N-1}$ (respectively $\{\Delta\varepsilon_{i+\frac{1}{2}} f_{i+\frac{1}{2}}\}_{1\leq i\leq N-1}$) and $\widetilde{\mathcal{M}}$ is the $(N-1)\times(N-1)$ tridiagonal matrix with components

$$\widetilde{\mathcal{M}}_{1,2} \,=\, -(\varepsilon_{\frac{5}{2}}^{\frac{3}{2}} - \varepsilon_{\frac{3}{2}}^{\frac{3}{2}})^{-1}\,,\;\; \widetilde{\mathcal{M}}_{1,1} \,=\, -\widetilde{\mathcal{M}}_{1,2} + (\varepsilon_{\frac{3}{2}}^{\frac{3}{2}})^{-1}$$

$$\widetilde{\mathcal{M}}_{i,i-1} \,=\, -(\varepsilon_{i+\frac{3}{2}}^{\frac{3}{2}} - \varepsilon_{i+\frac{1}{2}}^{\frac{3}{2}})^{-1}\,,\;\; \widetilde{\mathcal{M}}_{i,i+1} \,=\, -(\varepsilon_{i+\frac{1}{2}}^{\frac{3}{2}} - \varepsilon_{i-\frac{1}{2}}^{\frac{3}{2}})^{-1}\,,\;\; \widetilde{\mathcal{M}}_{i,i} \,=\, -\,(\widetilde{\mathcal{M}}_{i,i-1} + \widetilde{\mathcal{M}}_{i,i+1})\,,$$

$$\widetilde{\mathcal{M}}_{N-\frac{1}{2},N-\frac{3}{2}} \,=\, -\,(\varepsilon_{N-\frac{1}{2}}^{\frac{3}{2}} - \varepsilon_{N-\frac{3}{2}}^{\frac{3}{2}})^{-1}\,,\;\; \widetilde{\mathcal{M}}_{N-\frac{1}{2},N-\frac{1}{2}} \,=\, -\,\widetilde{\mathcal{M}}_{N-\frac{1}{2},N-\frac{3}{2}} - \frac{f_N}{\widetilde{E}_{N-\frac{1}{2}}}\,.$$

Proof. As in the previous proof we easily check that

$$\frac{\widetilde{D}_{i+\frac{3}{2}} - \widetilde{D}_{i+\frac{1}{2}}}{\varepsilon_{i+\frac{3}{2}}^{\frac{3}{2}} - \varepsilon_{i+\frac{1}{2}}^{\frac{3}{2}}} - \frac{\widetilde{D}_{i+\frac{1}{2}} - \widetilde{D}_{i-\frac{1}{2}}}{\varepsilon_{i+\frac{1}{2}}^{\frac{3}{2}} - \varepsilon_{i-\frac{1}{2}}^{\frac{3}{2}}} \,=\, -\,f_{i+\frac{1}{2}}\,\Delta\varepsilon_{i+\frac{1}{2}}\,,\quad 2\leq i\leq N-2\,. \tag{3.52}$$

We look what's happening at $i = N-1$. From the discretization of the diffusion term we get

$$\widetilde{D}_{N-\frac{1}{2}} \,=\, \sum_{j=1}^{N-1} \varepsilon_{j+\frac{1}{2}}^{\frac{3}{2}} f_{j+\frac{1}{2}}\Delta\varepsilon_{j+\frac{1}{2}} \,+\, \varepsilon_{N-\frac{1}{2}}^{\frac{3}{2}}\, f_{N+\frac{1}{2}}\,\Delta\varepsilon_{N+\frac{1}{2}}\,,$$

and

$$\widetilde{D}_{N-\frac{3}{2}} \,=\, \sum_{j=1}^{N-2} \varepsilon_{j+\frac{1}{2}}^{\frac{3}{2}} f_{j+\frac{1}{2}}\Delta\varepsilon_{j+\frac{1}{2}} \,+\, \varepsilon_{N-\frac{3}{2}}^{\frac{3}{2}} f_{N-\frac{1}{2}}\,\Delta\varepsilon_{N-\frac{1}{2}} \,+\, \varepsilon_{N-\frac{3}{2}}^{\frac{3}{2}} f_{N+\frac{1}{2}}\Delta\varepsilon_{N+\frac{1}{2}}\,.$$

Therefore

$$\frac{\widetilde{D}_{N-\frac{1}{2}} - \widetilde{D}_{N-\frac{3}{2}}}{\varepsilon^{\frac{3}{2}}_{N-\frac{1}{2}} - \varepsilon^{\frac{3}{2}}_{N-\frac{3}{2}}} = f_{N-\frac{1}{2}}\,\Delta\varepsilon_{N-\frac{1}{2}} + f_{N+\frac{1}{2}}\,\Delta\varepsilon_{N+\frac{1}{2}}\,.$$

But following (3.21) and (3.45) we can write $f_{N+\frac{1}{2}}\Delta\varepsilon_{N+\frac{1}{2}} = f_N\,\widetilde{D}_{N-\frac{1}{2}}/\widetilde{E}_{N-\frac{1}{2}}$, thus

$$\widetilde{D}_{N-\frac{1}{2}}\Big(\frac{1}{\varepsilon^{\frac{3}{2}}_{N-\frac{1}{2}} - \varepsilon^{\frac{3}{2}}_{N-\frac{3}{2}}} - \frac{f_N}{\widetilde{E}_{N-\frac{1}{2}}}\Big) - \frac{\widetilde{D}_{N-\frac{3}{2}}}{\varepsilon^{\frac{3}{2}}_{N-\frac{1}{2}} - \varepsilon^{\frac{3}{2}}_{N-\frac{3}{2}}} = f_{N-\frac{1}{2}}\,\Delta\varepsilon_{N-\frac{1}{2}}\,. \tag{3.53}$$

By summing the $(N-3)$ relations (3.52) together with (3.53) we easily get the equation at $i=1$. This ends the proof. □

Note that $\widetilde{\mathcal{M}}$ is a M-matrix if and only if $f_N < \widetilde{E}_{N-\frac{1}{2}}(\varepsilon^{\frac{3}{2}}_{N-\frac{1}{2}} - \varepsilon^{\frac{3}{2}}_{N-\frac{3}{2}})^{-1}$ and $f_N\varepsilon^{\frac{3}{2}}_{N-\frac{1}{2}}/\widetilde{E}_{N-\frac{1}{2}} \neq 1$.

3.1.5 Mass and energy conservations

Proposition 5. *The schemes $\mathcal{S}_1$ and $\mathcal{S}_2$ are conservative in density and energy.*

These conservation properties can be easily checked by taking $\{\phi_i\}_{1\le i\le N} = \{(1,\varepsilon_i)^t\}_{1\le i\le N}$ in (3.5) and (3.17) respectively.

3.1.6 Equilibrium solution

Recall that $M = \exp(-\beta\varepsilon)$ is the Maxwellian equilibrium solution of the equation (2.1) that has the same density and energy as $f(\varepsilon,0) = f_0(\varepsilon)$. We note $\{M_i\}_{1\le i\le N} = \{\exp(-\beta\varepsilon_i)\}_{1\le i\le N}$ the approximation of $\{M(\varepsilon_i)\}_{1\le i\le N}$ on the energy grid. In this section, we focus on the thermodynamical equilibrium state. We show in particular that the schemes $\mathcal{S}_1$ and $\mathcal{S}_2$ preserve the equilibrium state when is reached; it means that the numerical fluxes are equal to zero when the distribution f is equal to the equilibrium solution. First we consider the scheme $\mathcal{S}_1$.

Proposition 6. *We assume that $M_{i+\frac{1}{2}}$ is computed by Chang and Cooper average formula (3.11) where*

$$\delta_{i+\frac{1}{2}} = \frac{1}{\beta\Delta\varepsilon_{i+\frac{1}{2}}} - \frac{1}{\exp(\beta\Delta\varepsilon_{i+\frac{1}{2}})-1} \quad \forall i\,;\ 1\le i\le N-1\,. \tag{3.54}$$

Thus we have $M_{i+\frac{1}{2}} = -\dfrac{M_{i+1}-M_i}{\beta\Delta\varepsilon_{i+\frac{1}{2}}}$ and $\dfrac{E_{i+\frac{1}{2}}}{D_{i+\frac{1}{2}}} = \beta$ for all i such that $1\le i\le N-1$ where $E_{i+\frac{1}{2}}$ and $D_{i+\frac{1}{2}}$ are computed by the scheme $\mathcal{S}_1$ (see (3.9)).

Proof. First we remark that $M_{i+1} = \exp(-\beta\Delta\varepsilon_{i+\frac{1}{2}})M_i$ thus we have

$$M_{i+1} - M_i = -M_{i+1}\,(\exp(-\beta\Delta\varepsilon_{i+\frac{1}{2}}) - 1)\,,$$

Now from (3.11) and (3.54) we get

$$M_{i+\frac{1}{2}} \;=\; \left(\frac{1}{\beta\Delta\varepsilon_{i+\frac{1}{2}}} \;-\; \frac{1}{\exp(-\beta\Delta\varepsilon_{i+\frac{1}{2}})-1} \right) (M_i \;-\; M_{i+1}) \;+\; M_{i+1}\,.$$

Therefore

$$M_{i+\frac{1}{2}} \;=\; (\frac{1}{\beta\Delta\varepsilon_{i+\frac{1}{2}}} \;+\; \frac{M_{i+1}}{M_{i+1}-M_i})\,(M_i \;-\; M_{i+1}) \;+\; M_{i+1}\,,$$

and

$$M_{i+\frac{1}{2}} \;=\; -\,\frac{M_{i+1} \;-\; M_i}{\beta\Delta\varepsilon_{i+\frac{1}{2}}}\,. \tag{3.55}$$

Thanks to (3.55) the drift term reads at equilibrium

$$E_{i+\frac{1}{2}} = -\sum_{j=1}^{N-1} g_{i+\frac{1}{2},j+\frac{1}{2}}(M_{j+1}-M_j) = \sum_{j=1}^{N-1} g_{i+\frac{1}{2},j+\frac{1}{2}} M_{j+\frac{1}{2}} \beta\Delta\varepsilon_{j+\frac{1}{2}} = \beta D_{i+\frac{1}{2}}\,.$$

That completes the proof. □

Note that the Proposition 6 amounts to say that problem (3.13) has a solution at equilibrium given by relation (3.54). We are now going to show that at equilibrium this solution is unique. By the Brouwer fixed point theorem we saw that $\mathcal{H}$ has a fixed point in $[0,1]^{N-1}$. To prove the uniqueness of solution we just have to show that the mapping $\mathcal{H}$ is contracting, i.e.,$\| \mathcal{H}(x)-\mathcal{H}(y) \| \leq L' \| x-y \|$ for some norm, Lipschitz constant $L'<1$, and all x, y in $[0,1]^{N-1}$. Indeed, if $\mathcal{H}$ has two fixed points δ and δ' we get $\| \mathcal{H}(\delta)-\mathcal{H}(\delta') \| < \| \delta-\delta' \|$ and that is conflicting. We introduce the matrix $\boldsymbol{a}(\delta)=(\boldsymbol{a}_{i,j})=(\frac{\partial\mathcal{H}_{i+\frac{1}{2}}(\delta)}{\partial\delta_{j+\frac{1}{2}}})$ and the matrix norm induced by the maximum norm $\| \boldsymbol{a} \| = \max_i \sum_{j=1}^{N-1} | \boldsymbol{a}_{i,j} |$. At equilibrium we have

$$\boldsymbol{a}_{i,j} \;=\; -h'(\alpha_{i+\frac{1}{2}})\,\frac{E_{i+\frac{1}{2}}}{D^2_{i+\frac{1}{2}}} \Delta\varepsilon_{i+\frac{1}{2}}\, g_{i+\frac{1}{2},j+\frac{1}{2}} (M_j - M_{j+1}) \Delta\varepsilon_{j+\frac{1}{2}}\,.$$

Note that as M is a Maxwellian, M is monotone thus $M_j - M_{j+1} \geq 0$ and $E_{i+\frac{1}{2}} \geq 0$. Therefore

$$\sum_{j=1}^{N-1} | \boldsymbol{a}_{i,j} | = | h'(\alpha_{i+\frac{1}{2}}) |\, \frac{E_{i+\frac{1}{2}}}{D^2_{i+\frac{1}{2}}} \Delta\varepsilon_{i+\frac{1}{2}} \sum_{j=1}^{N-1} g_{i+\frac{1}{2},j+\frac{1}{2}} (M_j - M_{j+1}) \Delta\varepsilon_{j+\frac{1}{2}}\,.$$

We denote $\Delta\varepsilon = \max_i(\Delta\varepsilon_{i+\frac{1}{2}})$ and we get

$$\sum_{j=1}^{N-1} | \boldsymbol{a}_{i,j} | \leq | h'(\alpha_{i+\frac{1}{2}}) |\; \alpha^2_{i+\frac{1}{2}}\, \frac{\Delta\varepsilon}{\Delta\varepsilon_{i+\frac{1}{2}}}\,.$$

Now let $g(x) = h'(x)x^2$. By calculating g' it's easy to see that the function g is decaying over $[0,+\infty[$ and $g(0)=0$, $g(+\infty)=-1$. Therefore, since x is bounded we have $| g(x) | < 1$. By definition $\alpha_{i+\frac{1}{2}}$ reads

$$\alpha_{i+\frac{1}{2}} \;=\; \frac{\sum_{j=1}^{N-1} g_{i+\frac{1}{2},j+\frac{1}{2}} M_{j+1} (\exp(\beta\Delta\varepsilon_{j+\frac{1}{2}})-1)}{\sum_{j=1}^{N-1} g_{i+\frac{1}{2},j+\frac{1}{2}} ((1-\delta_{j+\frac{1}{2}})M_{j+1} + \delta_{j+\frac{1}{2}} M_j)}\, \Delta\varepsilon_{i+\frac{1}{2}}\,.$$

Because $E_{i+\frac{1}{2}}$ and $D_{i+\frac{1}{2}}$ are positive, $\delta_{i+\frac{1}{2}}$ lies in $[0, 1/2]$. Hence

$$\alpha_{i+\frac{1}{2}} \;\leq\; 2\frac{\sum_{j=1}^{N-1} g_{i+\frac{1}{2},j+\frac{1}{2}} M_{j+1}(\exp(\beta\Delta\varepsilon_{j+\frac{1}{2}})-1)}{\sum_{j=1}^{N-1} g_{i+\frac{1}{2},j+\frac{1}{2}} M_{j+1}\Delta\varepsilon_{j+\frac{1}{2}}}\Delta\varepsilon_{i+\frac{1}{2}}\,.$$

But $(\exp(x)-1)/x$ is an increasing function, consequently we have

$$\alpha_{i+\frac{1}{2}} \;\leq\; 2\frac{\exp(\beta\Delta\varepsilon)-1}{\Delta\varepsilon}\,\Delta\varepsilon_{i+\frac{1}{2}}\,,$$

and $\mid g(\alpha_{i+\frac{1}{2}}) \mid < 1$. It is clear that if $\{\Delta\varepsilon_{i+\frac{1}{2}}\}_{1\leq i\leq N}$ is a uniform sequence then $\sum_{j=1}^{N-1} \mid \boldsymbol{a}_{i,j} \mid < 1$ and $\mathcal{H}$ is contracting. Now if we admit that β is sufficiently small and $\Delta\varepsilon/\Delta\varepsilon_{i+\frac{1}{2}}$ close to the unity such that $\mid g(\alpha_{i+\frac{1}{2}})\Delta\varepsilon/\Delta\varepsilon_{i+\frac{1}{2}} \mid < 1$ then $\mathcal{H}$ is still contracting.

We turn now to scheme the $\mathcal{S}_2$.

Proposition 7. *We assume that $M_{i+\frac{1}{2}}$ is computed by the Chang and Cooper average formula (3.23) where*

$$\widetilde{\delta}_{i+\frac{1}{2}} \;=\; \frac{1}{\beta\Delta\varepsilon_{i+\frac{1}{2}}} \;-\; \frac{1}{\exp(\beta\Delta\varepsilon_{i+\frac{1}{2}})-1} \quad \forall i\,;\; 1\leq i\leq N-1\,.$$

Thus we have $\dfrac{\widetilde{E}_{i+\frac{1}{2}}}{\widetilde{D}_{i+\frac{1}{2}}} = \beta$ for all i such that $1\leq i\leq N$ where $\widetilde{E}_{i+\frac{1}{2}}$ and $\widetilde{D}_{i+\frac{1}{2}}$ are calculated with the scheme $\mathcal{S}_2$.

Proof. According to (3.29) and thanks to (3.21) we can write

$$\frac{\widetilde{D}_{N+\frac{1}{2}}}{\widetilde{E}_{N+\frac{1}{2}}} \;=\; \frac{\widetilde{D}_{N+\frac{1}{2}}}{\widetilde{E}_{N+\frac{1}{2}}}\,\frac{D_{N+\frac{1}{2}}\dfrac{\widetilde{E}_{N+\frac{1}{2}}}{\widetilde{D}_{N+\frac{1}{2}}} \;+\; g_{N+\frac{1}{2},N+\frac{1}{2}}\,f_N}{E_{N+\frac{1}{2}} \;+\; g_{N+\frac{1}{2},N+\frac{1}{2}}\,f_N}\,.$$

Therefore

$$\frac{\widetilde{D}_{N+\frac{1}{2}}}{\widetilde{E}_{N+\frac{1}{2}}} = \frac{D_{N+\frac{1}{2}}}{E_{N+\frac{1}{2}}} = \frac{1}{\beta}. \tag{3.56}$$

Once again we get

$$\frac{\widetilde{D}_{i+\frac{1}{2}}}{\widetilde{E}_{i+\frac{1}{2}}} \;=\; \frac{\widetilde{D}_{N+\frac{1}{2}}}{\widetilde{E}_{N+\frac{1}{2}}}\,\frac{D_{i+\frac{1}{2}}\dfrac{\widetilde{E}_{N+\frac{1}{2}}}{\widetilde{D}_{N+\frac{1}{2}}} \;+\; g_{i+\frac{1}{2},N+\frac{1}{2}}\,f_N}{E_{i+\frac{1}{2}} \;+\; g_{i+\frac{1}{2},N+\frac{1}{2}}\,f_N}\,,$$

and consequently $\dfrac{\widetilde{D}_{i+\frac{1}{2}}}{\widetilde{E}_{i+\frac{1}{2}}} = \dfrac{D_{i+\frac{1}{2}}}{E_{i+\frac{1}{2}}} = \dfrac{1}{\beta}$. This ends the proof. □

3.1.7 Positivity

If equation (2.1) is linear (in the sense that $E(f)$ and $D(f)$ are independent of f), Chang and Cooper [10] have shown that their scheme is positive; in other words, if initial value $f_0(\varepsilon)$ is non-negative then the value of $f(\varepsilon, t)$ at each successive time step remains non-negative. Afterwards, Larsen *et al.* [25] have pointed out that if (2.1) is non-linear (in particular if there exists non-linearity in energy variable) the Chang and Cooper method generally loses the property that numerical solutions are guaranteed to be positive. We are interested in the case where $E(f)$ and $D(f)$ depend on the distribution function. We show that $\mathcal{S}_1$ is a positive scheme while $\mathcal{S}_2$ can produce negative solution. Now we focus on the scheme $\mathcal{S}_1$.

Proposition 8. *We assume that for all distribution function $f = \{f_i\}_{1\leq i\leq N}$ there exists $\delta(f) = \{\delta_{i+\frac{1}{2}}(f)\}_{1\leq i\leq N-1}$ that is C^1 and given by (3.12). Then (3.7) has a solution. Furthermore, if $f_i(t=0)$ is non-negative, then $f_i(t)$ is non-negative for $t > 0$.*

Proof. In the first place as $\delta(f)$ is C^1 we are able to use Cauchy-Lipschitz's theorem. So that (3.7) has a unique solution. From the Proposition 1 we get $K_{i+\frac{1}{2}} = A_{i+\frac{1}{2}} f_{i+1} - B_{i+\frac{1}{2}} f_i$, with

$$A_{i+\frac{1}{2}} = v(\alpha_{i+\frac{1}{2}}) \frac{D_{i+\frac{1}{2}}}{\Delta\varepsilon_{i+\frac{1}{2}}} \quad \text{and} \quad B_{i+\frac{1}{2}} = u(\alpha_{i+\frac{1}{2}}) \frac{D_{i+\frac{1}{2}}}{\Delta\varepsilon_{i+\frac{1}{2}}}$$

and where

$$u(\alpha) = \frac{\alpha}{\exp(\alpha) - 1} \quad \text{and} \quad v(\alpha) = \frac{\alpha \exp(\alpha)}{\exp(\alpha) - 1}.$$

The function $u(x)$ is positive, decreasing from $-\infty$ to 0, while $v(x)$ is positive too, increasing from 0 to $+\infty$. To prove the positivity we use Lemma 1. Thus all we need to do is showing that the $A_{i+\frac{1}{2}}$'s and $B_{i+\frac{1}{2}}$'s are positive and bounded. As a result (see Remark 6) we have $D_{i+\frac{1}{2}} \geq 0$. Moreover u and v are positive functions therefore $A_{i+\frac{1}{2}}$ and $B_{i+\frac{1}{2}}$ are positive too. Now we have to show that $A_{i+\frac{1}{2}}$ and $B_{i+\frac{1}{2}}$ are bounded. At first we assume that f is a distribution function that has density $\overline{\boldsymbol{\rho}}$ and energy $\overline{\boldsymbol{\rho}\mathbf{E}}$ thus we get $f_j \leq \dfrac{\overline{\boldsymbol{\rho}}}{\max_i c_i}$, $\forall j$; $1 \leq j \leq N-1$. So that we can bound the drift term as following

$$\mid E_{i+\frac{1}{2}} \mid = \mid \sum_{j=1}^{N-1} g_{i+\frac{1}{2},j+\frac{1}{2}} (f_j - f_{j+1}) \mid \leq \frac{\overline{\boldsymbol{\rho}}}{\max_i c_i} 2 \sum_{j=1}^{N-1} g_{i+\frac{1}{2},j+\frac{1}{2}} \leq \frac{\overline{\boldsymbol{\rho}}}{\max_i c_i} 2(N-1) g_{N-\frac{1}{2},N-\frac{1}{2}} \,.$$

Consequently, there exists $C > 0$ such that $\mid E_{i+\frac{1}{2}} \mid < C$. In the same way we get an upperbound for the diffusion coefficients

$$D_{i+\frac{1}{2}} \leq \frac{\overline{\boldsymbol{\rho}}}{\max_i c_i} 2 \sum_{j=1}^{N-1} g_{i+\frac{1}{2},j+\frac{1}{2}} \Delta\varepsilon_{j+\frac{1}{2}} \,.$$

To continue we distinguish two cases. First we suppose that $\mid \alpha_{i+\frac{1}{2}} \mid \geq 1$. Then we get

$$A_{i+\frac{1}{2}} = \frac{\exp(\alpha_{i+\frac{1}{2}})}{\exp(\alpha_{i+\frac{1}{2}}) - 1} E_{i+\frac{1}{2}} \quad \text{and} \quad B_{i+\frac{1}{2}} = \frac{1}{\exp(\alpha_{i+\frac{1}{2}}) - 1} E_{i+\frac{1}{2}} \,.$$

And even if $\mid \alpha_{i+\frac{1}{2}} \mid \to \infty$, $A_{i+\frac{1}{2}}$ and $B_{i+\frac{1}{2}}$ remain bounded. Now we assume that $\mid \alpha_{i+\frac{1}{2}} \mid \leq 1$ then $u(\alpha_{i+\frac{1}{2}})$ and $v(\alpha_{i+\frac{1}{2}})$ are bounded and $A_{i+\frac{1}{2}}$ and $B_{i+\frac{1}{2}}$ too since $D_{i+\frac{1}{2}}$ is also bounded. Thus, thanks to Lemma 1 this ends the proof. □

Concerning the scheme $\mathcal{S}_2$ we have shown that the discretized diffusion term can become negative. This loss of positivity prevent us to provide an efficient result for the positivity of the scheme considering that our proof is based on the $\widetilde{D}_{i+\frac{1}{2}}$'s positivity. All we can do is supplying an example of contrary in the case where potentials are Maxwellian. Let f a distribution function that as density $\overline{\boldsymbol{\rho}} = \sum_{i=1}^{N} c_i f_i$ and energy $\overline{\boldsymbol{\rho}\mathbf{E}} = \sum_{i=1}^{N} c_i \varepsilon_i f_i$. We assume that $f = a\delta_{\varepsilon_1} + b\delta_{\varepsilon_N}$ where a and b are constants to define and the δ_{ε_i}'s are Dirac functions. We assure also that the integration domain is sufficient to check that $(\frac{3}{2}\mathbf{T} = \frac{\overline{\boldsymbol{\rho}\mathbf{E}}}{\overline{\boldsymbol{\rho}}} < \varepsilon_N = \mathcal{E})$. Now we have $\overline{\boldsymbol{\rho}} = c_1 a + c_N b$ and $\overline{\boldsymbol{\rho}\mathbf{E}} = c_N \varepsilon_N b$ (because $\varepsilon_1 = 0$). Thus $b = \overline{\boldsymbol{\rho}\mathbf{E}}/(c_N\varepsilon_N)$ and $a = (\overline{\boldsymbol{\rho}} - \overline{\boldsymbol{\rho}\mathbf{E}}/\varepsilon_N)/c_1$. Note that a and b are non-negative. From (3.19) we have

$$c_{N-1}\frac{df_{N-1}}{dt} \;=\; \widetilde{K}_{N-\frac{1}{2}} - \widetilde{K}_{N-\frac{3}{2}} = \widetilde{A}_{N-\frac{1}{2}} f_N - \widetilde{B}_{N-\frac{1}{2}} f_{N-1} - (\widetilde{A}_{N-\frac{3}{2}} f_{N-1} - \widetilde{B}_{N-\frac{3}{2}} f_{N-2})\,,$$

therefore $\frac{df_{N-1}}{dt} = \widetilde{A}_{N-\frac{1}{2}} b/c_{N-1}$. Now thanks to (3.28), $\widetilde{A}_{N-\frac{1}{2}}$ and $\widetilde{D}_{N-\frac{1}{2}}$ have the same sign. Taking again what we saw previously (equation (3.33)) we write

$$\frac{\widetilde{D}_{N-\frac{1}{2}}}{\varepsilon_{N-\frac{1}{2}}^{\frac{3}{2}}}(1 \;-\; \varepsilon_{N+\frac{1}{2}}^{\frac{3}{2}} \frac{f_N}{\frac{3}{2}\overline{\boldsymbol{\rho}}}) \;=\; \sum_{j=1}^{N-1} \varepsilon_{j+\frac{1}{2}}^{\frac{3}{2}}\, f_{j+\frac{1}{2}}\, \Delta\varepsilon_{j+\frac{1}{2}}\,.$$

Subsequently $1 - \varepsilon_{N+\frac{1}{2}}^{\frac{3}{2}} \frac{f_N}{\frac{3}{2}\overline{\boldsymbol{\rho}}} = 1 - \frac{\overline{\boldsymbol{\rho}\mathbf{E}}}{\frac{3}{2}\overline{\boldsymbol{\rho}}} \frac{\varepsilon_{N+\frac{1}{2}}^{\frac{3}{2}}}{c_N\varepsilon_N} = 1 - \overline{\mathbf{T}}\frac{\varepsilon_{N+\frac{1}{2}}^{\frac{3}{2}}}{\varepsilon_N}\frac{1}{\sqrt{\varepsilon_N}\Delta\varepsilon_N}$. If $\Delta\varepsilon_N \longrightarrow 0$ it is clear that $\widetilde{D}_{N-\frac{1}{2}}$ becomes negative. Therefore $\frac{df_{N-1}}{dt} < 0$ and consequently f_{N-1} gets negative too. For the Coulombian case we can take the same example of contrary to show the non-positivity of the scheme $\mathcal{S}_2$.

We can be more precise with $\mathcal{S}_2$ in the case where potentials are Maxwellian. Let us recall that we denoted $Y = \frac{3}{2}\overline{\boldsymbol{\rho}} - \varepsilon_{N+\frac{1}{2}}^{\frac{3}{2}} f_N$ (see formula (3.34)). As usual, f is a non-negative distribution that has non-zero density $\overline{\boldsymbol{\rho}}$ and energy $\overline{\boldsymbol{\rho}\mathbf{E}}$ such that $\overline{\mathbf{T}} = 2/3\frac{\overline{\boldsymbol{\rho}\mathbf{E}}}{\overline{\boldsymbol{\rho}}} < T_{\max}$. These functions generate a compact subset of $\mathbb{R}^N$: $\mathcal{D} = \{f \in \mathbb{R}^N;\ \forall i,\ \ f_i \geq 0, \sum_i c_i f_i = \overline{\boldsymbol{\rho}}$ and $\sum_i c_i\varepsilon_i f_i = \overline{\boldsymbol{\rho}\mathbf{E}}\}$. Now we show that $\widetilde{D}_{i+\frac{1}{2}}$ behaves like $1/Y$. First, we saw that the drift coefficients are always positive and the diffusion coefficients (namely $\widetilde{D}_{i+\frac{1}{2}} = \varepsilon_{i+\frac{1}{2}}^{\frac{3}{2}} d(f)$) have all the same sign. Consequently the Chang and Cooper coefficients always verify: $0 \leq \delta_{i+\frac{1}{2}} \leq \frac{1}{2}$ if $Y > 0$ and $\frac{1}{2} \leq \delta_{i+\frac{1}{2}} \leq 1$ if $Y < 0$. To begin, we assume that $Y > 0$. Thus, thanks to (3.33) we get the inequality

$$Y\widetilde{D}_{N-\frac{1}{2}}/(\frac{3}{2}\overline{\boldsymbol{\rho}}) = \varepsilon_{N-\frac{1}{2}}^{\frac{3}{2}}\widetilde{d}(f)(1-\frac{\varepsilon_{N+\frac{1}{2}}^{\frac{3}{2}} f_N}{\frac{3}{2}\overline{\boldsymbol{\rho}}}) = \sum_{j=1}^{N-1} g_{N-\frac{1}{2},j+\frac{1}{2}} f_{j+\frac{1}{2}}\Delta\varepsilon_{j+\frac{1}{2}} \geq \sum_{j=1}^{N-1} g_{N-\frac{1}{2},j+\frac{1}{2}} f_{j+1}\Delta\varepsilon_{j+\frac{1}{2}}.$$

The functional $\sum_{j=1}^{N-1} g_{N-\frac{1}{2},j+\frac{1}{2}} f_{j+1}\Delta\varepsilon_{j+\frac{1}{2}}$ is linear continuous on $\mathcal{D}$ thus achieves its minimum on this compact subset. This minimum is strictly positive because if it is equal to zero, that implies that all the mass will be concentrated at ε_1, consequently the energy would be

zero and we have supposed that the energy is strictly positive.
On the other hand, if $Y < 0$, (3.33) gives

$$Y\widetilde{D}_{N-\frac{1}{2}}/(\frac{3}{2}\overline{\boldsymbol{\rho}}) \geq \sum_{j=1}^{N-1} g_{N-\frac{1}{2},j+\frac{1}{2}} f_j \Delta\varepsilon_{j+\frac{1}{2}},$$

and the right-hand side admits a strictly positive minimum (if not, all the mass will be concentrated at ε_N and we have supposed that we only consider distribution function such that $0 \leq \overline{\mathbf{T}} \leq T_{\max}$).
This shows that for the diffusion coefficient $\widetilde{D}_{N-\frac{1}{2}}$ we ge $\lim_{Y\to\pm 0} \widetilde{D}_{N-\frac{1}{2}} \to \pm\infty$. Since all the diffusion coefficient are proportionals and have the same sign, one has for all i: $\lim_{Y\to\pm 0} \widetilde{D}_{i+\frac{1}{2}} \to \pm\infty$.
In other words $\widetilde{D}_{i+\frac{1}{2}}$ behaves like $\dfrac{1}{Y}$. This result says also that the scheme $\mathcal{S}_2$ is not consistent with the truncated FPL equation defined by (3.1) but only with (2.3).
In the Coulombian case, the situation is more complicated. But we are able to construct a sequence of positive distribution functions f^ν, with constant mass and energy, which have f^0 as limit as $\nu \to \pm 0$ and such that:

- f^0 verifies $Y = 0$, where Y is given by (3.48),
- there exist an index i_0 and a constant a such that $\min(f_{i_0}, f_{i_0+1}) \geq a > 0$,
- $f^0_{N-1} \geq b f^0_N$, with $b > 2$.

We begin by the construction of f^0. First we assume that f^0 lies in $\mathcal{D}$ and we set

$$\frac{1}{b} f^0_{N-1} = f^0_N = \varepsilon^{\frac{3}{2}}_{N+\frac{1}{2}}/\frac{3}{2}\overline{\boldsymbol{\rho}}.$$

The mass and the energy due to these two points are equal respectively to

$$\Delta\overline{\boldsymbol{\rho}} = \varepsilon^{\frac{3}{2}}_{N+\frac{1}{2}}(bc_{N-1} + c_N)/\frac{3}{2}\overline{\boldsymbol{\rho}} \quad \text{and} \quad \Delta\overline{\boldsymbol{\rho}\mathbf{E}} = \varepsilon^{\frac{3}{2}}_{N+\frac{1}{2}}(b\varepsilon_{N-1}c_{N-1} + \varepsilon_N c_N)/\frac{3}{2}\overline{\boldsymbol{\rho}}.$$

Since $c_i \simeq \Delta\varepsilon_i\sqrt{\varepsilon_i}$, for $\Delta\varepsilon_N$ and $\Delta\varepsilon_{N-1}$ sufficiently small we have $\overline{\boldsymbol{\rho}} - \Delta\overline{\boldsymbol{\rho}} > 0$ and $\overline{\boldsymbol{\rho}\mathbf{E}} - \Delta\overline{\boldsymbol{\rho}\mathbf{E}} > 0$. Then we choose on the $N-2$ first points a positive distribution with mass $\overline{\boldsymbol{\rho}} - \Delta\overline{\boldsymbol{\rho}}$ and with energy $\overline{\boldsymbol{\rho}\mathbf{E}} - \Delta\overline{\boldsymbol{\rho}\mathbf{E}}$, by example a Maxwellian having this mass and this energy. The resulting distribution f^0 verifies all the three points above.

The functional $Y = \frac{3}{2}\overline{\boldsymbol{\rho}} - \varepsilon^{\frac{3}{2}}_{N+\frac{1}{2}} f_N$ is continuous on the compact set $\mathcal{D}$ and $\{0\} \in \mathcal{I}m(\overset{o}{Y}(\mathcal{D}))$, since for the maxwellian that has density $\overline{\boldsymbol{\rho}}$ and energy $\overline{\boldsymbol{\rho}\mathbf{E}}$, $Y > 0$ and for the example of contrary above $Y < 0$. Thus there exist sequences $f^\nu \in \mathcal{D}$ such that $\lim_{\nu\to 0} f^\nu = f^0$ with $Y^\nu \geq 0$ or $Y^\nu \leq 0$ and with $Y^\nu \to 0$. For ν sufficiently small we have, by example, $\min(f^\nu_{i_0}, f^\nu_{i_0+1}) \geq a/2$ and $f^0_{N-1} > b/2b f^0_N$. For this sequence $Y^\nu > 0$ we use the definition of the diffusion coefficients

$$Y^\nu \widetilde{D}^\nu_{N-\frac{1}{2}} \geq \frac{a}{2} \sum_j g_{N-\frac{1}{2},j+\frac{1}{2}} \Delta\varepsilon_{j+\frac{1}{2}} \frac{3}{2}\overline{\boldsymbol{\rho}}.$$

Thus if $Y^\nu \geq 0$, $\lim_{\nu\to 0} \widetilde{D}^\nu_{N-\frac{1}{2}} = +\infty$ or if $Y^\nu \leq 0$, $\lim_{\nu\to 0} \widetilde{D}^\nu_{N-\frac{1}{2}} = -\infty$.

Since the definition of Y is the same in the Coulombian and in the Maxwellian cases, the above example works also in the Maxwellian case. We have seen that in the domain $Y > 0$ the diffusion coefficients are positive. And naturally the question that arise is: is this domain stable by (3.19)? it seems that this is not true: for the sequence f^ν constructed above one can verify easily using (3.26) that $\frac{df^\nu_N}{dt}$ tends to $+\infty$ as Y^ν tends to $+0$ that is $\frac{dY^\nu}{dt}$ tends to $-\infty$ as Y^ν tends to $+0$. This suggests that one can attain the boundary $Y = 0$ in finite time. And for $Y = 0$, as we have seen, the scheme $\mathcal{S}_2$ is not defined.

Despite its deficiencies for the positivity one could use the scheme $\mathcal{S}_2$: a solution to avoid the lack of positivity would be to take the domain of computation sufficiently large in order to have the tail of the distribution function sufficiently small and thus $Y \simeq \frac{3}{2}\overline{\boldsymbol{\rho}}$.

3.1.8 Discussion: some remarks and about Epperlein's version of the Chang and Cooper method.

We begin by given some remarks about the above analysis.

Remark 14. *It is easy to check that the above results about positivity, energy conservation and equilibrium states for $\mathcal{S}_1$ and $\mathcal{S}_2$ don't depend on the definition of the coefficients c_i, see (3.3), or on our choice for the value of ε_1.*

Remark 15. *For $\mathcal{S}_1$ (and $\mathcal{S}_2$ if the scheme is positive) if one takes the Chang and Cooper coefficients $\delta_{i+\frac{1}{2}}$ at equilibrium, that is given by (3.54), or freeze the value given by the initial data, it is not difficult to see that if it is done in the expression of the fluxes given by (3.10) the scheme is still positive but no more conservative in energy and if it is done in (3.14) the scheme is still conservative in energy but no more positive. The algebra that permits to pass from the form (3.10) to the form (3.14) is valid if and only if the coefficients verify at all times the relation (3.13). This remark is still valid if one does not calculate exactly the Chang and Cooper coefficients by an iterative method.*

Remark 16. *The scheme $\mathcal{S}_2$ is in fact not consistent with the truncated FPL equation (3.1), but only with (2.1).*

Remark 17. *We don't have any H-theorem for $\mathcal{S}_1$ (and $\mathcal{S}_2$) except for $\mathcal{S}_1$ in the case where potentials are maxwellian and the energy grid is uniform, since in this case the scheme reduces to the entropic scheme of Berezin-Pekker.*

Remark 18. *The scheme $\mathcal{S}_2$ is not positive but it would not be a surprise if implicit time discretization leads to positive solutions. Actually it is well known that implicit time discretization adds some kind of numerical "viscosity" which "stabilizes" algorithms.*

We are also interesting in the Epperlein velocity space discretization for solving the FPLE for Coulombian collisions [22] since it is based on the method of Chang and Cooper. This method is used in the code SPARK developed by Epperlein [21] or by Kingham and Bell in the recent code IMPACT [23]. As in the present paper the collisions coefficients are defined by Langdon. To difference the equation in velocity space, the author applies the Chang and Cooper approach. In his paper Epperlein develops a fully implicit finite-difference method

for solving FPL equation and claims that his approach conserves both energy and number density exactly. Without talking about time discretization we just focus on his velocity space discretization and properties. In particular the Epperlein scheme differs on our scheme only by boundary conditions at $\varepsilon = \mathcal{E}$. Actually, the author assumes that the diffusion term and the distribution function vanish at $\varepsilon_{N+\frac{1}{2}}$ and that the distribution function vanishes at ε_{N+1}. Thus the numerical flux is zero and consequently the density is conserved. In fact the Epperlein velocity discretization is nothing else than a miscellany of what we called scheme $\mathcal{S}_1$ and scheme $\mathcal{S}_2$: the drift term is the one discretized by scheme $\mathcal{S}_2$, formula (3.39), and the diffusion term the one computed by $\mathcal{S}_1$, formula (3.37). Now, we can prove that the Epperlein velocity space discretization produces non-negative solution in consideration that the diffusion term is non-negative. Whereas we raise a doubt concerning the energy conservation. Effectively, assuming that diffusion vanishes at boundary while drift not, breaks the symmetry in the right-hand side of the semi-discretized equation. And the energy conservation proof leans on the symmetry between ε and ε' (or v and v'). At last (3.56) fails and this relation is indispensable to check that the Chang and Cooper type scheme preserves the equilibrium state when is reached. As for $\mathcal{S}_2$ this scheme is not consistent with the truncated FPL equation (3.1), but only with (2.1).

Remark 19. *We can remark that these three Chang and Cooper type schemes are equivalent when the integrating domain is not bounded. The only difference between these schemes lies in the manner to treat boundary conditions when the integrating domain is reduced to* $[0, \mathcal{E}]$.

3.2 Alternative schemes

The Chang and Cooper method is not the only way to provide positive, conservative and equilibrium states preserving schemes for the FPL equation. As we have seen above it is a very complicated scheme for this equation, not always positive or conservative in energy, depending on the boundary condition taken at the end of the domain of computation. The schemes we'll propose in the next sections share also these properties and are simpler.

3.2.1 Equilibrium scheme (scheme $\mathcal{S}_3$)

The first scheme is based on the work of Larsen *et al.* [25]. Their work is for linear and non-linear (in the sense that collisions terms are non-linear in energy variables) Fokker-Planck equations. But they do not consider non-linearity as in the Landau equation, that is drift and diffusion coefficients are functionnals of the distribution functions. One of the two main ideas exposed in their paper to preserve Maxwellian states is to remark that ε-derivative that appears in the flux can be rewritten as

$$\frac{\partial}{\partial \varepsilon} = -\beta y \frac{\partial}{\partial y}$$

where $y = \exp(-\beta\varepsilon)$. Thus it becomes easy to preserve Maxwellians in the weak form of the Fokker-Planck equation. Following these authors, in the case of the FPL equation, we set

$$\overline{\Delta\varepsilon}_{i+\frac{1}{2}} = -\frac{(\Delta M)_{i+\frac{1}{2}}}{\beta M_{i+1}} = \frac{\exp(\beta\Delta\varepsilon_{i+\frac{1}{2}}) - 1}{\beta} \simeq \Delta\varepsilon_{i+\frac{1}{2}} \tag{3.57}$$

Thus we consider the following approximation of the weak symmetrized form of the problem

$$\sum_{i=1}^{N} c_i \frac{\partial f_i}{\partial t}\phi_i = -\frac{1}{2}\sum_{i=1}^{N-1}\sum_{j=1}^{N-1}\left(\frac{\Delta\phi_{i+\frac{1}{2}}}{\Delta\varepsilon_{i+\frac{1}{2}}} - \frac{\Delta\phi_{j+\frac{1}{2}}}{\Delta\varepsilon_{j+\frac{1}{2}}}\right) g_{i+\frac{1}{2},j+\frac{1}{2}}\left(f_{j+\frac{1}{2}}\frac{\Delta f_{i+\frac{1}{2}}}{\overline{\Delta\varepsilon}_{i+\frac{1}{2}}} - f_{i+\frac{1}{2}}\frac{\Delta f_{j+\frac{1}{2}}}{\overline{\Delta\varepsilon}_{j+\frac{1}{2}}}\right)\overline{\Delta\varepsilon}_{j+\frac{1}{2}}\overline{\Delta\varepsilon}_{i+\frac{1}{2}}, \tag{3.58}$$

with the approximations $f_{i+\frac{1}{2}}$ taken decentered and defined by $f_{i+\frac{1}{2}} = f_{i+1}$.
As in the previous section we obtain the system of ordinary equation for the approximation of the FPL equation

$$\frac{df_i}{dt} = Q_i^{\mathcal{S}_3} \quad 1 \le i \le N \tag{3.59}$$

where $Q_1^{\mathcal{S}_3} = K_{\frac{3}{2}}/c_1$, $Q_i^{\mathcal{S}_3} = (K_{i+\frac{1}{2}} - K_{i-\frac{1}{2}})/c_i$ for $2 \le i \le N-1$ and $Q_N^{\mathcal{S}_3} = -K_{N-\frac{1}{2}}/c_N$ with the numerical flux

$$K_{i+\frac{1}{2}} = \frac{\overline{\Delta\varepsilon}_{i+\frac{1}{2}}}{\Delta\varepsilon_{i+\frac{1}{2}}}\sum_{j=1}^{N-1} g_{i+\frac{1}{2},j+\frac{1}{2}}(f_{j+1}\frac{\Delta f_{i+\frac{1}{2}}}{\overline{\Delta\varepsilon}_{i+\frac{1}{2}}} - f_{i+1}\frac{\Delta f_{j+\frac{1}{2}}}{\overline{\Delta\varepsilon}_{j+\frac{1}{2}}})\overline{\Delta\varepsilon}_{j+\frac{1}{2}} \text{ for } i; 1 \le i \le N-1$$

.
By factorizing the terms f_i in the last sum, the numerical flux reads

$$K_{i+\frac{1}{2}} = \frac{\overline{\Delta\varepsilon}_{i+\frac{1}{2}}}{\Delta\varepsilon_{i+\frac{1}{2}}}(\frac{\Delta f_{i+\frac{1}{2}}}{\overline{\Delta\varepsilon}_{i+\frac{1}{2}}}\sum_{j=1}^{N-1} g_{i+\frac{1}{2},j+\frac{1}{2}} f_{j+1}\overline{\Delta\varepsilon}_{j+\frac{1}{2}} + f_{i+1}\sum_{j=1}^{N-1}(g_{i+\frac{1}{2},j+\frac{1}{2}} - g_{i+\frac{1}{2},j-\frac{1}{2}})f_j$$
$$- f_{i+1}f_N g_{i+\frac{1}{2},N-\frac{1}{2}}).$$

If we bring together the terms f_{i+1} we get

$$K_{i+\frac{1}{2}} = \frac{\overline{\Delta\varepsilon}_{i+\frac{1}{2}}}{\Delta\varepsilon_{i+\frac{1}{2}}}((\frac{1}{\overline{\Delta\varepsilon}_{i+\frac{1}{2}}}\sum_{j=1}^{N-1} g_{i+\frac{1}{2},j+\frac{1}{2}} f_{j+1}\overline{\Delta\varepsilon}_{j+\frac{1}{2}} + \sum_{j=1}^{N-1}(g_{i+\frac{1}{2},j+\frac{1}{2}} - g_{i+\frac{1}{2},j-\frac{1}{2}})f_j - f_N g_{i+\frac{1}{2},N-\frac{1}{2}})f_{i+1}$$
$$- \sum_{j=1}^{N-1} g_{i+\frac{1}{2},j+\frac{1}{2}} f_{j+1}\overline{\Delta\varepsilon}_{j+\frac{1}{2}}\frac{f_i}{\overline{\Delta\varepsilon}_{i+\frac{1}{2}}}).$$

We denote

$$A_{i+\frac{1}{2}} = \frac{\overline{\Delta\varepsilon}_{i+\frac{1}{2}}}{\Delta\varepsilon_{i+\frac{1}{2}}}\left(\frac{1}{\overline{\Delta\varepsilon}_{i+\frac{1}{2}}}\sum_{j=1}^{N-1} g_{i+\frac{1}{2},j+\frac{1}{2}} f_{j+1}\overline{\Delta\varepsilon}_{j+\frac{1}{2}} + \sum_{j=1}^{N-1}(g_{i+\frac{1}{2},j+\frac{1}{2}} - g_{i+\frac{1}{2},j-\frac{1}{2}})f_j - f_N g_{i+\frac{1}{2},N-\frac{1}{2}}\right), \tag{3.60}$$

and

$$B_{i+\frac{1}{2}} = \frac{1}{\Delta\varepsilon_{i+\frac{1}{2}}}\sum_{j=1}^{N-1} g_{i+\frac{1}{2},j+\frac{1}{2}} f_{j+1}\overline{\Delta\varepsilon}_{j+\frac{1}{2}}, \tag{3.61}$$

thus we have $K_{i+\frac{1}{2}} = A_{i+\frac{1}{2}}f_{i+1} - B_{i+\frac{1}{2}}f_i$.
We can summarize the properties of this scheme in the following proposition

Proposition 9. *If $g_{i+\frac{1}{2},j+\frac{1}{2}}$ is an increasing sequence and $\Delta\varepsilon_{i+\frac{1}{2}} \le \Delta\varepsilon_{N-\frac{1}{2}}$ the scheme $\mathcal{S}_3$ is conservative in mass and energy. Moreover it is positive and preserves the equilibrium state when is reached.*

Proof. Mass and energy conservation are easily checked by setting $\phi = 1$ and $\phi = \varepsilon$ in the weak symmetrized form of the discrete FPL equation (3.58).
Using the definition (3.57) of $\overline{\Delta\varepsilon}_{i+\frac{1}{2}}$ and the definition of the $f_{i+\frac{1}{2}}$'s it's easy to see that for $f = M$ we have

$$f_{j+\frac{1}{2}}\frac{\Delta f_{i+\frac{1}{2}}}{\overline{\Delta\varepsilon}_{i+\frac{1}{2}}} - f_{i+\frac{1}{2}}\frac{\Delta f_{j+\frac{1}{2}}}{\overline{\Delta\varepsilon}_{j+\frac{1}{2}}} = \beta(M_{j+1}M_{i+1} - M_{i+1}M_{j+1}) = 0$$

thus the scheme preserves the Maxwellians states.
Let us now check the positivity. Let us recall the numerical flux associated to (3.59)

$$K_{i+\frac{1}{2}} = A_{i+\frac{1}{2}}f_{i+1} - B_{i+\frac{1}{2}}f_i,$$

where $A_{i+\frac{1}{2}}$ and $B_{i+\frac{1}{2}}$ are given respectively by (3.60) and (3.61). Now, to use Lemma 1 we have to check that $A_{i+\frac{1}{2}}$ and $B_{i+\frac{1}{2}}$ are positive and bounded. As $g_{i+\frac{1}{2},j+\frac{1}{2}}$ is positive, $B_{i+\frac{1}{2}}$ is obviously positive. We can write

$$A_{i+\frac{1}{2}} = \frac{\overline{\Delta\varepsilon}_{i+\frac{1}{2}}}{\Delta\varepsilon_{i+\frac{1}{2}}}\Big(\frac{1}{\overline{\Delta\varepsilon}_{i+\frac{1}{2}}}\sum_{j=1}^{N-2} g_{i+\frac{1}{2},j+\frac{1}{2}}f_{j+1}\overline{\Delta\varepsilon}_{j+\frac{1}{2}} + \sum_{j=1}^{N-1}(g_{i+\frac{1}{2},j+\frac{1}{2}} - g_{i+\frac{1}{2},j-\frac{1}{2}})f_j$$
$$+ f_N g_{i+\frac{1}{2},N-\frac{1}{2}}\Big(\frac{\overline{\Delta\varepsilon}_{N-\frac{1}{2}}}{\overline{\Delta\varepsilon}_{i+\frac{1}{2}}} - 1\Big)\Big).$$

Assuming that $g_{i+\frac{1}{2},j+\frac{1}{2}}$ is an increasing sequence and that $\Delta\varepsilon_{i+\frac{1}{2}} \leq \Delta\varepsilon_{N-\frac{1}{2}}$ leads to the positivity of $A_{i+\frac{1}{2}}$. Now, $\overline{\Delta\varepsilon}_{i+\frac{1}{2}}/\Delta\varepsilon_{i+\frac{1}{2}} \in [0,1]$ therefore $A_{i+\frac{1}{2}}$, $B_{i+\frac{1}{2}}$ are bounded since, due to mass conservation, we have $f_i \leq \dfrac{\overline{\boldsymbol{\rho}}}{\min_j c_j}$. According to Lemma 1, f_i cannot vanish in finite time. That completes the proof. □

Remark 20. *One could replace the exact equilibrium M_i in the definition of the scheme by an approximate equilibrium state $\widetilde{M}_i$ with same mass and energy and the resulting scheme is still conservative, positive but the equilibrium state is now $\widetilde{M}_i$. This could be useful since this scheme requires the knowledge of the equilibrium state and normally this can be done only by solving a non-linear equation with an iterative method.*

Remark 21. *If we stand $f_{i+\frac{1}{2}} = f_i$ instead of $f_{i+\frac{1}{2}} = f_{i+1}$ in the numerical flux and if we take $\overline{\Delta\varepsilon}_{i+\frac{1}{2}} = \dfrac{(\Delta M)_{i+\frac{1}{2}}}{\beta M_i}$, the Maxwellians are also preserved but we can show that $\mathcal{S}_3$ is positive if and only if the energy grid is uniform.*

Remark 22. *On a uniform grid all the terms $\overline{\Delta\varepsilon}_{i+\frac{1}{2}}$ are equal and up to a multiplicative constant the flux reduces to*

$$K_{i+\frac{1}{2}} = \sum_{j=1}^{N-1} g_{i+\frac{1}{2},j+\frac{1}{2}}(f_{i+1}f_j - f_i f_{j+1}).$$

The scheme is then nothing else than the scheme provided by Berezin and Pekker [1] which is also an entropic scheme as shown in [5].

Remark 23. *For Maxwellian ($g(\varepsilon,\varepsilon') = \varepsilon^{\frac{3}{2}}\varepsilon'^{\frac{3}{2}}$) or Coulombian ($g(\varepsilon,\varepsilon') = \min(\varepsilon^{\frac{3}{2}},\varepsilon'^{\frac{3}{2}})$) potentials the evaluation of all the coefficients $A_{i+\frac{1}{2}}$ and $B_{i+\frac{1}{2}}$ can be achieved in only $\mathcal{O}(N)$ operations, as explained in [4, 5].*

Remark 24. *The second idea of Larsen et al. in [25] to preserve Maxwellians and applied here to the isotropic FPL equation is to write $\dfrac{\partial f}{\partial \varepsilon} = M\dfrac{\partial}{\partial \varepsilon}(\dfrac{f}{M}) - \beta f$. The isotropic FPL equation reads*

$$\int_0^\infty \frac{\partial f(\varepsilon)}{\partial t}\phi(\varepsilon)\sqrt{\varepsilon}d\varepsilon = \\ -\frac{1}{2}\int_0^\infty\int_0^\infty (\frac{\partial\phi(\varepsilon)}{\partial\varepsilon} - \frac{\partial\phi(\varepsilon')}{\partial\varepsilon'})g(\varepsilon,\varepsilon')\left(f(\varepsilon')M(\varepsilon)\frac{\partial}{\partial\varepsilon}(\frac{f(\varepsilon)}{M(\varepsilon)}) - f(\varepsilon)M(\varepsilon')\frac{\partial}{\partial\varepsilon'}(\frac{f(\varepsilon')}{M(\varepsilon')})\right)d\varepsilon' d\varepsilon.$$

Proceeding as above leads to a scheme which is indeed conservative in mass and energy and preserves the Maxwellians. But at this time, we cannot show the positivity. Nevertheless it could be possible to derive a positive scheme since there are sufficiently degrees of freedom in the discretization, namely the factor $f_{i+\frac{1}{2}}$ and $M_{i+\frac{1}{2}}$.

3.2.2 Entropy decaying scheme (scheme $\mathcal{S}_4$)

The present approach consists in deriving a discretization from the "Log" form (2.4) because this helps to check easily the main properties of the operator, conservation and H-theorem which are given without proof. This approach was initiated by Degond and Lucquin in [17] for the 3-D Landau equation (for the implementation see [7, 26]), used for the 2-D axisymetric Landau equation [26], and applied in the isotropic case in [3, 5]. We present a new version of the algorithm based on the "Log" form.

From the previous section we retain the approximation of $\displaystyle\int_0^{\varepsilon_0}\frac{\partial f}{\partial t}\phi\sqrt{\varepsilon}d\varepsilon$ as $\displaystyle\sum_{i=1}^{N} c_i\frac{df_i}{dt}\phi_i$. We now turn to the discretization of the right-hand side of (2.4)

$$(r.h.s.) = -\frac{1}{2}\sum_{i=1}^{N-1}\sum_{j=1}^{N-1}\left(\int_{\varepsilon_i}^{\varepsilon_{i+1}}\int_{\varepsilon_j}^{\varepsilon_{j+1}}(\frac{\partial\phi(\varepsilon)}{\partial\varepsilon} - \frac{\partial\phi(\varepsilon')}{\partial\varepsilon'})g(\varepsilon,\varepsilon')f(\varepsilon')f(\varepsilon)(\frac{\partial\log f(\varepsilon)}{\partial\varepsilon} - \frac{\partial\log f(\varepsilon')}{\partial\varepsilon'})d\varepsilon d\varepsilon'\right). \tag{3.62}$$

For each integrals of (3.62) we use again a midpoint quadrature formula, the ε-derivative are approximated by centered finite difference operator, thus

$$\begin{aligned}(r.h.s.) = &-\frac{1}{2}\sum_{i=1}^{N-1}\sum_{j=1}^{N-1}\left(\frac{\Delta\phi_{i+\frac{1}{2}}}{\Delta\varepsilon_{i+\frac{1}{2}}} - \frac{\Delta\phi_{j+\frac{1}{2}}}{\Delta\varepsilon_{j+\frac{1}{2}}}\right)g_{i+\frac{1}{2},j+\frac{1}{2}}k_{i+\frac{1}{2},j+\frac{1}{2}} \\ &\left(\frac{(\Delta\log f)_{i+\frac{1}{2}}}{\Delta\varepsilon_{i+\frac{1}{2}}} - \frac{(\Delta\log f)_{j+\frac{1}{2}}}{\Delta\varepsilon_{j+\frac{1}{2}}}\right)\Delta\varepsilon_{i+\frac{1}{2}}\Delta\varepsilon_{j+\frac{1}{2}},\end{aligned} \tag{3.63}$$

with $g_{i+\frac{1}{2},j+\frac{1}{2}} = g(\varepsilon_{i+\frac{1}{2}},\varepsilon_{j+\frac{1}{2}})$ and the terms $k_{i+\frac{1}{2},j+\frac{1}{2}}$ stand for a second-order approximation of $f_{i+\frac{1}{2}}f_{j+\frac{1}{2}}$ at the center of the interval $[\varepsilon_i,\varepsilon_{i+1}]\times[\varepsilon_j,\varepsilon_{j+1}]$. In the paper of Berezin *et al.* [1], the terms $k_{i+\frac{1}{2},j+\frac{1}{2}}$ are of the form $k_{i+\frac{1}{2},j+\frac{1}{2}} = k_{i+\frac{1}{2}}k_{j+\frac{1}{2}}$ where the $k_{i+\frac{1}{2}}$ are taken as

an arithmetic mean of f_i and f_{i+1}. This yields a discrete model for which it cannot be proved the f remains positive as it must be. More recently, Cordier and Buet [4] have chosen a second-order approximation as the harmonic average; that is $(2f_if_{i+1})/(f_i+f_{i+1})$. They proved the existence of a unique, positive and entropy solution for the semi-discretized FPL equation. Unfortunately, it turns out if initial data is not strictly positive. Furthermore, if the distribution function is zero at a point all numerical fluxes vanish. In the case of a uniform grid, in [5], and from an idea of [6], the authors have considered the formula

$$k_{i+\frac{1}{2},j+\frac{1}{2}} = \frac{f_i(\Delta f)_{j+\frac{1}{2}} - f_j(\Delta f)_{i+\frac{1}{2}}}{(\Delta \log f)_{j+\frac{1}{2}} - (\Delta \log f)_{i+\frac{1}{2}}} \quad \text{if} \quad (\Delta \log f)_{i+\frac{1}{2}} \neq (\Delta \log f)_{j+\frac{1}{2}},$$

$$\text{and} \quad k_{i+\frac{1}{2},j+\frac{1}{2}} = f_if_j \qquad \text{when} \quad (\Delta \log f)_{i+\frac{1}{2}} = (\Delta \log f)_{j+\frac{1}{2}}.$$

For a uniform grid and only in this case the above expression can be simplified into

$$k_{i+\frac{1}{2},j+\frac{1}{2}} = \frac{f_if_{j+1} - f_jf_{i+1}}{\log(f_{j+1}f_i) - \log(f_{i+1}f_j)},$$

with such assumption they recover the scheme proposed in [1]. Following this approach we take

$$k_{i+\frac{1}{2},j+\frac{1}{2}} = l_{i+\frac{1}{2}}l_{j+\frac{1}{2}} \left(\frac{f_{i+1}^{a_{i+\frac{1}{2}}} f_j^{a_{j+\frac{1}{2}}} - f_i^{a_{i+\frac{1}{2}}} f_{j+1}^{a_{j+\frac{1}{2}}}}{\log(f_{i+1}^{a_{i+\frac{1}{2}}} f_j^{a_{j+\frac{1}{2}}}) - \log(f_i^{a_{i+\frac{1}{2}}} f_{j+1}^{a_{j+\frac{1}{2}}})} \right),$$

with $a_{i+\frac{1}{2}} = \dfrac{\Delta\varepsilon}{\Delta\varepsilon_{i+\frac{1}{2}}}$, where $\Delta\varepsilon = \max_j \Delta\varepsilon_{j+\frac{1}{2}}$ and for consistency we take $l_{i+\frac{1}{2}} = (\dfrac{f_i+f_{i+1}}{2})^{1-a_{i+\frac{1}{2}}}$.
By the relation $(x-y)(\log x - \log y) \geq 0$ we assure that $k_{i+\frac{1}{2},j+\frac{1}{2}}$ remains positive. Thus we obtain the following weak formulation for the semi-discretized model

$$\sum_{i=1}^{N} c_i \frac{df_i}{dt}\phi_i = -\sum_{i=1}^{N-1}(\phi_{i+1}-\phi_i) \sum_{j=1}^{N-1} g_{i+\frac{1}{2},j+\frac{1}{2}} (\frac{f_{i+1}+f_i}{2})^{1-a_{i+\frac{1}{2}}} (\frac{f_{j+1}+f_j}{2})^{1-a_{j+\frac{1}{2}}}$$
$$\left(f_{i+1}^{a_{i+\frac{1}{2}}} f_j^{a_{j+\frac{1}{2}}} - f_i^{a_{i+\frac{1}{2}}} f_{j+1}^{a_{j+\frac{1}{2}}} \right) \frac{\Delta\varepsilon_{j+\frac{1}{2}}}{\Delta\varepsilon}.$$

As in the previous sections the system of ordinary equation reads

$$\frac{df_i}{dt} = Q_i^{\mathcal{S}_4}, \quad 1 \leq i \leq N \tag{3.64}$$

where $Q_1^{\mathcal{S}_4} = K_{\frac{3}{2}}/c_1$, $Q_i^{\mathcal{S}_4} = (K_{i+\frac{1}{2}} - K_{i-\frac{1}{2}})/c_i$ for $2 \leq i \leq N-1$ and $Q_N^{\mathcal{S}_4} = -K_{N-\frac{1}{2}}/c_N$ with the numerical flux

$$K_{i+\frac{1}{2}} = \sum_{j=1}^{N-1} g_{i+\frac{1}{2},j+\frac{1}{2}} (\frac{f_{i+1}+f_i}{2})^{1-a_{i+\frac{1}{2}}} (\frac{f_{j+1}+f_j}{2})^{1-a_{j+\frac{1}{2}}} \left(f_{i+1}^{a_{i+\frac{1}{2}}} f_j^{a_{j+\frac{1}{2}}} - f_i^{a_{i+\frac{1}{2}}} f_{j+1}^{a_{j+\frac{1}{2}}} \right) \frac{\Delta\varepsilon_{j+\frac{1}{2}}}{\Delta\varepsilon}.$$

To simplify we denote $\boldsymbol{\Theta}^r_{i+\frac{1}{2}} = (2f_{i+1}/(f_i+f_{i+1}))^{a_{i+\frac{1}{2}}-1}$ and $\boldsymbol{\Theta}^l_{i+\frac{1}{2}} = (2f_i/(f_i+f_{i+1}))^{a_{i+\frac{1}{2}}-1}$ this provides

$$K_{i+\frac{1}{2}} = \sum_{j=1}^{N-1} g_{i+\frac{1}{2},j+\frac{1}{2}} (\boldsymbol{\Theta}^r_{i+\frac{1}{2}} f_{i+1} \boldsymbol{\Theta}^l_{j+\frac{1}{2}} f_j - \boldsymbol{\Theta}^l_{i+\frac{1}{2}} f_i \boldsymbol{\Theta}^r_{j+\frac{1}{2}} f_{j+1}) \frac{\Delta\varepsilon_{j+\frac{1}{2}}}{\Delta\varepsilon}.$$

We write

$$A_{i+\frac{1}{2}} = \mathbf{\Theta}^r_{i+\frac{1}{2}}(\sum_{j=1}^{N-1} g_{i+\frac{1}{2},j+\frac{1}{2}} \frac{\Delta\varepsilon_{j+\frac{1}{2}}}{\Delta\varepsilon} \mathbf{\Theta}^l_{j+\frac{1}{2}} f_j), \tag{3.65}$$

and

$$B_{i+\frac{1}{2}} = \mathbf{\Theta}^l_{i+\frac{1}{2}}(\sum_{j=1}^{N-1} g_{i+\frac{1}{2},j+\frac{1}{2}} \frac{\Delta\varepsilon_{j+\frac{1}{2}}}{\Delta\varepsilon} \mathbf{\Theta}^r_{j+\frac{1}{2}} f_{j+1}), \tag{3.66}$$

thus

$$K_{i+\frac{1}{2}} = A_{i+\frac{1}{2}} f_{i+1} - B_{i+\frac{1}{2}} f_i.$$

Proposition 10. *The scheme $\mathcal{S}_4$ is conservative in mass and energy and preserves the positivity and the Maxwellian equilibrium when it is reached. Moreover it is an entropic scheme.*

Proof. The conservation of mass and energy are directly checked on (3.63) by taking $\phi = 1$ or $\phi = \varepsilon$. Let M the Maxwellian equilibrium solution. The proof is straightforward by substituting the distribution function M also in (3.63). If we define the entropy by

$$H(f) = \sum_{i=1}^{N} c_i f_i \log f_i,$$

thus using (3.63) with $\phi = \log f$, one has

$$\frac{dH(f)}{dt} = \sum_{i=1}^{N} c_i \frac{df_i}{dt} \log f_i \leq 0,$$

and the equality occurs when $f = \exp(-\beta\varepsilon)$, that is when f is a Maxwellian. In other terms we have a H-theorem for this scheme. The numerical flux associated to scheme $\mathcal{S}_4$ is

$$K_{i+\frac{1}{2}} = A_{i+\frac{1}{2}} f_{i+1} - B_{i+\frac{1}{2}} f_i,$$

where $A_{i+\frac{1}{2}}$ and $B_{i+\frac{1}{2}}$ are respectively given by (3.65) and (3.66). If $\{f_i\}_{1\leq i\leq N}$ is positive it's straightforward to show that

$$0 \leq \mathbf{\Theta}^l_{i+\frac{1}{2}} = \left(2f_i/(f_{i+1}+f_i)\right)^{a_{i+\frac{1}{2}}-1} \leq 1 \quad \text{and} \quad 0 \leq \mathbf{\Theta}^r_{i+\frac{1}{2}} = \left(2f_{i+1}/(f_{i+1}+f_i)\right)^{a_{i+\frac{1}{2}}-1} \leq 1$$

thus $A_{i+\frac{1}{2}}$ and $B_{i+\frac{1}{2}}$ are positive and upper bounded since, due to mass conservation, we have $f_i \leq \dfrac{\overline{\boldsymbol{\rho}}}{\min_j c_j}$. Thus according to Lemma 1 the solution cannot vanish in finite time and consequently it is positive. That ends the proof. □

Remark 25. *On a uniform grid the flux reduces to*

$$K_{i+\frac{1}{2}} = \sum_{j=1}^{N-1} g_{i+\frac{1}{2},j+\frac{1}{2}} (f_{i+1} f_j - f_i f_{j+1}),$$

thus the scheme reduces to the scheme of Berezin and Pekker [1] which is also an entropic scheme as shown in [5].

Remark 26. *For Maxwellian ($g(\varepsilon,\varepsilon') = \varepsilon^{\frac{3}{2}}\varepsilon'^{\frac{3}{2}}$) or Coulombian ($g(\varepsilon,\varepsilon') = \min(\varepsilon^{\frac{3}{2}},\varepsilon'^{\frac{3}{2}})$) potentials the evaluation of all the coefficients $A_{i+\frac{1}{2}}$ and $B_{i+\frac{1}{2}}$ can be achieved in only $\mathcal{O}(N)$ operations as explained in [4, 5].*

Remark 27. *For consistency the coefficients $a_{i+\frac{1}{2}}$ must remain bounded as $\Delta\varepsilon \to 0$.*

Remark 28. *One could choose the coefficients $g_{i+\frac{1}{2}}$ as $(\frac{f_i + f_{i+1} + \mu}{2})^{1-a_{i+\frac{1}{2}}}$ where $\mu << 1$ is a threshold to avoid undefined coefficients $\mathbf{\Theta}^r_{i+\frac{1}{2}}$ $\mathbf{\Theta}^l_{i+\frac{1}{2}}$ if $f_i = f_{i+1} = 0$ and that does not change anything in the properties of the scheme.*

Remark 29. *In a view of a time discretization it is easy to see that one can freeze the coefficients $\mathbf{\Theta}^r_{i+\frac{1}{2}}$ $\mathbf{\Theta}^l_{i+\frac{1}{2}}$, for example by taking their values at time zero, and the resulting scheme preserves mass, energy, positivity and Maxwellian state but is no more entropic.*

4 Conclusions

In this work, we have analyzed the discretization of the isotropic FPL equation in order to obtain positive, conservative and Maxwellian state preserving schemes.
Despite the lack of existence and uniqueness results for the schemes presented here, we have clarified the derivation as well as the properties of the Chang and Cooper method in this non-linear case. We have clearly explained how bad boundary conditions could lead to negative or non conservative schemes. As we have seen, the Chang and Cooper method is a quite complicated scheme for a non linear equation such as the FPL equation. We have also shown that the Epperlein scheme could not guarantee energy conservation and Maxwellian preserving, and the scheme $\mathcal{S}_2$ (developed by Langdon in the case of Coulombian potentials [15, 24]), is not a positive scheme. Only $\mathcal{S}_1$, the new scheme we developed, fulfills all the requirements. Let us recall that on infinite grids these three schemes are identical. The differences stem from the truncation of the domain of computation and the choice of the boundary conditions. In conclusion a direct discretization of the special form (2.15) and (2.16) of the drift and diffusion coefficients is not a guarantee to ensure energy conservation as claimed by Kingham and Bell in [23].
We have also presented and analyzed in detail two simpler schemes, namely $\mathcal{S}_3$ and $\mathcal{S}_4$ which are not based on the Chang and Cooper method. However, they are as efficient as $\mathcal{S}_1$. The scheme $\mathcal{S}_4$ based on the "Log" form of the operator is also an entropic scheme. Let us also mention that the Logarithm of the distribution function would not be a problem for an implicit time discretization since the solution of the FPL equation would be positive for all time $t > 0$ as soon as the initial data is non-negative.
We can summarize the properties of the above schemes as follow:

- **Chang and Cooper $\mathcal{S}_1$: positive, conserves the energy and the Maxwellians.**
- **Chang and Cooper $\mathcal{S}_2$: non-positive, conserves the energy and the Maxwellians.**
- **Chang and Cooper by Epperlein: positive, does not conserve the energy and the Maxwellians.**
- **$\mathcal{S}_3$: positive, conserves the energy and the Maxwellians.**

- $\mathcal{S}_4$**: positive, conserves the energy and the Maxwellians and is entropic.**

We hope that some of the remarks would help people to design an implicit time discretization both positive and conservative.
Concerning the implementation of implicit time discretization, our conviction is that a fixed point method based on the form of the fluxes $K_{i+\frac{1}{2}} = A_{i+\frac{1}{2}} f_{i+1} + B_{i+\frac{1}{2}} f_i$ by taking the coefficients $A_{i+\frac{1}{2}}, B_{i+\frac{1}{2}}$ at the previous iteration is sufficient. At each iteration there is no energy conservation (see also [22]) but the iteration is positive, and by making the fixed point converge to zero machine size to enforce energy conservation would be efficient enough. This is the method chosen in [8] in the context of granular media for the same type of equation and by Kingham and Bell in the code IMPACT for the isotropic FPL equation [23]. But people could also follow the alternative ideas developped by Lemou and Mieussens in [28] for implicit time discretization of the 3-D or isotropic Landau equation.

Acknowledgments

We wish to thank A. Decoster for many useful discussions during the course of this work.

References

[1] A. YU, BEREZIN, V.N. KHUDICK and M.S. PEKKER: *Conservative finite difference schemes for the Fokker-Planck equation not violating the law of an increasing entropy.* J. Comput. Phys., Vol. 69, pp. 163-174, 1987.

[2] A. V. BOBYLEV and V. A CHUYANOV: *On the numerical solution of Landau's kinetic equation.* USSR Comput Math. Math. Phys. 16, No 2, 1976.

[3] C. BUET and S. CORDIER: *Numerical analysis of conservatives and entropy schemes for the Fokker-Planck-Landau equation.* SIAM J. of Numer. Anal., Vol. 36, p. 953, 1999.

[4] C. BUET and S. CORDIER: *Conservative and entropy schemes for the isotropic Fokker-Planck-Landau equation.* J. Comput. Phys. 145, p. 228, 1998.

[5] C. BUET and S. CORDIER: *Numerical analysis of the isotropic Fokker-Planck-Landau equation.* J. Comput. Phys. 179, No 1, pp. 43-67, 2002.

[6] C. BUET, S. DELLACHERIE and R. SENTIS: *Numerical solution of an ionic Fokker-Planck equation with electronic temperature.* SIAM J. Numer. Anal. 39, No 4, pp. 1219-1253, 2001.

[7] C. BUET, S. CORDIER, P. DEGOND and M. LEMOU: *Fast algorithms for numerical, conservative, and entropy approximations of the Fokker-Planck-Landau equation.* J. Comput. Phys. 133, No 2, pp. 310-322, 1997.

[8] C. BUET, S. CORDIER and V. DOS SANTOS: *A Conservative and Entropy Scheme for a Simplified Model of Granular Media.* Transp. Theory Stat. Phys. 33, No 2, pp. 125-155, 2004.

[9] C. BUET and S. DELLACHERIE: *About the Chang and Cooper method for linear Fokker-Planck equations.* Unpublished.

[10] J. S. CHANG and G. COOPER: *A practical difference scheme for Fokker-Planck equations.* J. Comput. Phys., Vol. 6, Issue 1, August 1970.

[11] H. COHN: *Numerical integration of the Fokker-Planck equation and the evolution of stars clusters.* The Astrophysical Journal 234, pp. 1036-1053, 1979.

[12] H. COHN: *Late core collapse in star clusters and the gravothermal instability.* The Astrophysical Journal 242, pp. 765-771, 1980.

[13] M. CROUSEIX and A. MIGNOT: *Analyse numérique des équations différentielles.* Masson, 1983.

[14] A. DECOSTER: private communication.

[15] A. DECOSTER and A. B. LANGDON: unpublished 1981.

[16] P. DEGOND and B. LUCQUIN-DESREUX: *The Fokker-Planck asymptotics of the Boltzmann collision operator in the Coulomb case.* Math. Models and Methods in Appl.Sci, Vol. 2, No 2, pp. 167-182, 1992.

[17] P. DEGOND and B. LUCQUIN-DESREUX: *An entropy scheme for the Fokker-Planck collision operator of plasma kinetic theory.* Numer. Math. 68, No 2, pp.239-262, 1994.

[18] L. DESVILLETTES: *On Asymptotics of the Boltzmann Equation when the Collisions Become Grazing.* Transport Theory and Statistical Physics, Vol. 21, No 3, pp. 259-276, 1992.

[19] L. DESVILLETTES and C. VILLANI: *On the spatially homogeneous Landau equation for hard potentials. Part I: Existence, uniqueness and smoothness.* Comm. P.D.E 25, 1-2, pp.179-259, 2000.

[20] L. DESVILLETTES and C. VILLANI: *On the spatially homogeneous Landau equation for hard potentials. Part II: H-Theorem and applications.* Comm. P.D.E 25, 1-2, pp. 261-298, 2000.

[21] E. M. EPPERLEIN: *Fokker-Planck modelling of electrons transport in laser-produced plasmas.* Laser and particles Beams, Vol. 2, No 2, pp. 257-272, 1994.

[22] E. M. EPPERLEIN: *Implicit and conservative difference schemes for the Fokker-Planck equation.* J. Comput. Phys., Vol. 112, p. 291, 1994.

[23] R. J. KINGHAM and A.R. BELL: *An implicit Vlasov-Fokker-Planck code to model non-local electron transport in 2-D with magnetic fields.* J. Comput. Phys. 194, issue 1, pp. 1-34. 2004.

[24] A.B. LANGDON: *Conservative differencing of the electron Fokker-Planck transport equation.* CECAM Report of Workshop on The Flux Limiter and Heat Flow Instabilities in Laser-Fusion Plasmas, Universite Paris Sud, France, 1981.

[25] E. W. LARSEN, C. D. LEVERMORE, G. C. POMRANING and J. G. SANDERSON: *Discretization methods for one-dimensional Fokker-Planck operators.* J. Comput. Phys., Vol. 61, Issue 3, December 1985.

[26] M. LEMOU: *Multipole expansions for the Fokker-Planck-Landau operator.* Numer. Math. 78, No 4, pp. 597-618, 1998.

[27] M. LEMOU: *Numerical Algorithms for Axisymmetric Fokker-Planck-Landau Operators.* J. Comput. Phys., Vol. 157, Issue 2, 20 January 2000.

[28] M. LEMOU and L. MIEUSSENS: *Implicit schemes for the Fokker-Planck-Landau equation.* SIAM J. Sci. Comp. 27(3), pp. 809-830, 2005.

[29] L. SPITZER and R. HARM: *Transport phenomena in a completely ionized gaz.* Phys. rev, Vol. 89, pp. 977-981, 1953.

[30] I.P. SHKAROFSKY, T.W. JOHNSTON and M.P. BACHYNSKI: *The Particle Kinetics of Plasmas.* Addison Wesley, 1966.

ÉDITÉ PAR
LA DIRECTION DES SYSTEMES
D'INFORMATION

CEA / SACLAY 91191 GIF-SUR-YVETTE CEDEX FRANCE